

# Better Together? CEO Identity and Firm Productivity\*

Inês Black<sup>†</sup>

November 21, 2017

JOB MARKET PAPER  
[Click Here for Latest Version](#)

## Abstract

This paper analyzes the relationship between CEO quality and firm productivity in the private sector using CEO job transitions. Using a matched employer-employee data set, I attempt to disentangle the role of the CEO type (identity) from that of the firm in revenue productivity and evaluate the existence and relevance of match complementarities between CEO and firm types. I present a proxy measure of CEO quality that takes advantage of differential patterns of CEO mobility throughout their careers 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 estimated 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 indicate that 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-manager, Firm Productivity, CEO-Firm Match, Endogenous Mobility

---

\*I am very grateful to my advisor Caterina Calsamiglia and my sponsor at UCL, Imran Rasul, for their invaluable 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 for helpful comments and suggestions. I also acknowledge the support of the faculty and administrative staff at the Departments of Economics at UCL and Bank of Portugal's during my research visits in which I developed this paper. All mistakes are my own.

<sup>†</sup>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](mailto: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 top-management<sup>1</sup> 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<sup>2</sup>. 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<sup>3</sup>; 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 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<sup>4</sup>. 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

---

<sup>1</sup>I use the term "CEO" throughout the remainder of the paper. However, under the CEO label I include any and all head-manager of a firm with more than 20 employees.

<sup>2</sup>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.

<sup>3</sup>Gabaix & Landier (2008); Terviö (2008); Malmendier & Tate (2009); Bebchuk et al. (2011)

<sup>4</sup>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 their 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)<sup>5</sup>.

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)<sup>6</sup>. 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<sup>7,8</sup>. 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 a firm production model in which CEO and firm-type matches

---

<sup>5</sup>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.

<sup>6</sup>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.

<sup>7</sup>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.

<sup>8</sup>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.

are allowed to result in complementarities in production. This model relaxes the assumptions of the previous analysis by incorporating the 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 *kmeans* clustering algorithm<sup>9</sup> 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 this distributional approach estimates a finite mixture model which assumes that the impact of CEO and firm heterogeneity types in production are each a result of a mixture of Gaussian distributions that generate separate 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 one observed.

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<sup>10</sup>. Moreover, results regarding CEO and firm specific contributions to wage setting are in line with the literature<sup>11</sup>. 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 2% of average revenues, with a stronger effect (3%) in the top 10<sup>th</sup> percentile of the revenue distribution.

This paper contributes to the literature of organization economics and corporate governance, which has made significant progress in documenting a relationship between top-management<sup>12</sup>, in corporate outcomes<sup>13</sup>, either through their characteris-

---

<sup>9</sup>Steinley (2006).

<sup>10</sup>Pérez-González (2006); Bennedsen et al. (2007).

<sup>11</sup>Bender et al. (2016).

<sup>12</sup>Chief Executive Officers (CEO), Chief Financial Officers (CFO) and Chief Operating Officers (COO).

<sup>13</sup>Profits, ROA, ROI, M&A decisions among others.

tics (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-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. In this paper, I build on on these groups of studies by (i) addressing challenge of non-random assignment of managers to firms when analyzing panel data sets by putting forth a proxy measure of CEO quality that attenuates the endogenous CEO mobility, (ii) generalizing the scope of the analysis of the role of CEOs in the firm, which is typically placed on very large and/or publicly traded firms, to include all medium and large firms in the Portuguese private sector and (iii) documenting evidence regarding the importance of CEO-firm complementarities in production<sup>14</sup>.

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 (Terzio, 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. I contribute to this discussion by explicitly analyzing the role of the CEO-firm match complementarities in the firm's productivity. I find evidence that a non-negligible part firm revenue productivity can be attributed to the complementarities between the CEO and the firm.

This paper also fits in 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)<sup>15</sup>. 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<sup>16</sup>. 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. I use their environment to evaluate the role CEO-firm complementarities, under en-

---

<sup>14</sup>This issue is addressed also in a recent paper by Bandiera et al. (2017) who develop structural CEO-firm match model.

<sup>15</sup>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).

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

dogenous CEO mobility, in firm productivity.

The remainder of the paper is organized as follows. Section 2 details the theoretical 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<sup>17</sup>. 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<sup>18</sup>.

Let the described economy consist of  $J$  firms,  $j \in \{1, \dots, J\}$ , at each time-period  $t \in \{1, \dots, T\}$ . 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  $\Lambda : \mathbb{R}^+ \rightarrow [\underline{A}, \bar{A}]$ . Moreover, firms can hire from a pool of CEOs,  $i \in \{1, \dots, N\}$ . 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  $\Gamma : \mathbb{R}^+ \rightarrow [\underline{\alpha}, \bar{\alpha}]$  at each time  $t$ . CEO quality is assumed to be permanent and unidimensional (Becker, 1973). I also assume that CEOs are hired according to their success as employees was (Peter et al., 1969). Managerial ability ( $\alpha_i$ ) enters the production function in two ways. First, as an input (Bender et al., 2016). Second,

---

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

<sup>18</sup>An illustrative example would be to think of this quality as an identity type, something that is particular to the CEO.

through a decreasing returns to scale (DRS) transformation of the production function, reflecting the limited span of control of the CEO<sup>19</sup>.

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  $\alpha_i$  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:

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

where  $\alpha_i$  and  $A_j$  correspond to the CEO quality type and firm technology, respectively. Given the limited span of control of the CEO, 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:

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

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  $\alpha_i$ . Labor and capital can be hired at equilibrium prices  $w$  and  $r$ , respectively<sup>20</sup>.

For the remainder of the paper, I focus on gross revenue as the object of the firm's maximization problem<sup>21</sup>. That is,

$$PY_{j,t} = (P * A_j) \alpha_i^\mu L_{j,t}^\delta K_{j,t}^{1-\delta-\mu} \quad (3)$$

where  $P * A_j$  is TFPR (or revenue TFP<sup>22</sup>) 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.*

$$\frac{dPY_{j,t}}{d\alpha_i} > 0 \quad (4)$$

*Proof:* see Appendix A.

<sup>19</sup>The intuition being that the CEOs' management and supervisory abilities are inversely proportional to the quantity of L and K units under their control, *ceteris paribus*.

<sup>20</sup>I assume firm size is small enough not to change equilibrium prices.

<sup>21</sup>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.

<sup>22</sup>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 homogeneous capital. Let the described economy consist of the same  $J$  firms,  $j \in \{1, \dots, J\}$ , at each time-period  $t \in \{1, \dots, T\}$ . 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,  $i \in \{1, \dots, N\}$ . I assume CEO's quality is exogenous ( $\alpha_i$ ). 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  $\alpha_i$  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  $\delta$  represents the firm's production function with limited CEO span-of-control:

$$\ln(Y_{j,t}) = \theta_{i,j} + \mu \ln(\alpha_i) + \ln(A_j) + \delta \ln(L_{j,t}) + (1 - \delta - \mu) \ln(K_{j,t}) \quad (5)$$

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.*

$$\frac{dPY_{j,t}}{d\theta_{i,j}} > 0 \quad (6)$$

*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)<sup>23</sup>. The

<sup>23</sup>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%<sup>24</sup>. Despite the similarities in labor participation, labor productivity in Portugal is significantly lower than that of the EU<sup>25</sup>. 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 than 1-2% a year<sup>26</sup>. 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<sup>27</sup>. Most (68.2%) of these firms' employment is dedicated to services, comparable to a 72% in the EU<sup>28</sup>.

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<sup>29</sup> that have at least one employee on payroll during the survey reference week<sup>30</sup>. 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<sup>31</sup> and bargaining and unions<sup>32</sup>. The mandatory character of both *Quadros de Pessoal* and *IES*, together with reporting based on tax-authority valid profiles<sup>33</sup>, 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<sup>34</sup>. 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,

---

<sup>24</sup>Source: OECD (2014), Share of employed who are managers.

<sup>25</sup>Source: PORDATA and Eurostat (2016).

<sup>26</sup>Source: Bank of Portugal.

<sup>27</sup>Source: PORDATA (2017), *Empresas, Pequenas e Médias Empresas*.

<sup>28</sup>OECD (2015), Employment in the services sector.

<sup>29</sup>Public administration and informal market services are excluded. Includes private, nonprofit and public firms.

<sup>30</sup>One week of October of each year.

<sup>31</sup>Cardoso et al. (2016).

<sup>32</sup>Card et al. (2015); Addison et al. (2017).

<sup>33</sup>All employee salaries and employer sales revenues are reported as they are declared to the Portuguese Tax Authority (*Direção Geral dos Impostos*) and Social Security.

<sup>34</sup>The survey has waves after 2013; however, these are the ones available at the Bank of Portugal.

hours worked, overtime, and earnings<sup>35</sup>). 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<sup>36</sup>. 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%<sup>37</sup> winsorization of wage outliers<sup>38</sup>.

Table 1 presents means and standard deviations<sup>39</sup> for two samples: non-CEO and CEO employees. Column 1 presents the results for CEO employees. Female CEOs make up 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. Table 2 presents the same statistics for the largest connected set of non-CEO employees and firms<sup>40</sup>. 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 3 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<sup>41</sup>, Financial Manager, other high-level managers and mid-level managers. Other high-level managers consist of operative managers 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, Tables 1, 2 and 3 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.

---

<sup>35</sup>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.

<sup>36</sup>Appendix B elaborates on the methodology used to identify the firm CEO.

<sup>37</sup>I set all variables below the 1% percentile to the 1% percentile of the distribution; the same with 99% percentile.

<sup>38</sup>Appendix B provides further details on sample selection criteria.

<sup>39</sup>Summary statistics for Table 1 are constructed using the average of cross-section estimates.

<sup>40</sup>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.

<sup>41</sup>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 4 presents summary statistics on firm characteristics<sup>42</sup>. A share of approximately 28% of all firms is located in Lisbon, whereas 19% are in Porto. Approximately 37% of the firms operate in the manufacturing sector, 14% in construction and 49% in the service sector. 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  $i = 1, \dots, N$  employees and  $j = 1, \dots, J$  firms. I model the logarithm of wages as a function of employee observables, firm and employee fixed effects:

$$y_{it} = \alpha_i + \psi_{j(i,t)} + \mathbf{X}_{it}\boldsymbol{\beta} + \varepsilon_{it} \quad (7)$$

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,  $\psi_{j(i,t)}$  represents firm  $j$  fixed effect, which captures constant firm specific heterogeneity,  $\alpha_i$  are the employee fixed effects and  $\mathbf{X}_{it}$  represents a vector of time-varying employee-level control variables<sup>43</sup> and year fixed effects. The parameter of interest is  $\alpha_i$ , which I interpret as an innate ability that is valued in the labor market in the same way. I use  $\alpha_i$  as the estimated employee quality<sup>44</sup>. I recover the

---

<sup>42</sup>Summary statistics for Table 4 are made up of averages of annual cross-sections, except in the case of firm longevity.

<sup>43</sup>Quadratic terms in age fully interacted with schooling levels. More details on variables and coding are presented in Appendix B.

<sup>44</sup>A definition of “quality” in the context of this paper is further described in Appendix B of this paper.

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<sup>45</sup>.

I follow previous work<sup>46</sup> 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 similar summary statistics between the largest connected set and the full sample<sup>47</sup> (Table 2). [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 (7) derives from the job mobility<sup>48</sup> of employees across firms. This mobility 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 (7) in three parts, as [Card et al. \(2013\)](#), to highlight potential cases where the stated orthogonality condition does not hold. The error term  $\varepsilon_{it}$  is composed by a random employee-firm match effect,  $(\lambda_{j(i,t)})$ , a unit-root process that reflects increments in employee quality  $(\omega_i)$  and a transitory shock component  $(v_{j(i,t)})$  as described in equation (8):

$$\varepsilon_{it} = \lambda_{i,j(i,t)} + \omega_{it} + v_{it} \tag{8}$$

---

<sup>45</sup>As a counter example, in the case of an employee who stays in the same firm throughout the whole panel time-span, the employee and firm fixed effects cannot be disentangled.

<sup>46</sup>[Card et al. \(2013\)](#); [Cardoso et al. \(2016\)](#).

<sup>47</sup>See Appendix C for a more nuanced analysis of the differences in the largest connected set.

<sup>48</sup>I label a job mobility “event” by identifying employees whose associated firm identifier changes from one period (year) to the next.

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<sup>49</sup>. I apply the same logic in the context of this model, where match effects are reflected in wages. The unit root component  $\omega_{it}$  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  $\omega_{it}$  has mean zero for each employee in their observed time period. Finally, there is a transitory term ( $v_{it}$ ) 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  $\mathbf{y}$  denote the stacked  $NT \times 1$  vector of year-sorted employee wages (where  $NT = N$ ),  $\mathbf{E}$  denote the  $NT \times N$  design matrix of employee indicators,  $\mathbf{F}$  is the  $NT \times J$  design matrix of firm indicators and  $\mathbf{X}$  is a  $NT \times J$  matrix of time-varying employee covariates and  $\varepsilon$  denotes the error term. Equation (7) can be written in matrix notation as:

$$\mathbf{y} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{F}\boldsymbol{\psi} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (9)$$

Consistent estimation of equation (9) via OLS implies the following assumptions regarding the interaction of the error terms with explanatory variables:

$$E[e^i \varepsilon] = 0, \forall i \quad E[f^j \varepsilon] = 0, \forall j \quad E[x^k \varepsilon] = 0, \forall k \quad (10)$$

Equation (10) describes OLS orthogonality conditions between the error term  $\varepsilon$  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<sup>50</sup>. 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<sup>51</sup>. If this type of endogenous mobility is present in the CEO data, causal identification of equation (9) 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 (9) boils down to the verification of the assumption  $E[f^j \varepsilon] = 0, \forall j$ <sup>52</sup>. Appendix C provides intuition for a simple example with two CEOs, two firms and two

<sup>49</sup>Mortensen & Pissarides (1994); Shimer & Smith (2000); Postel-Vinay & Robin (2002); Eeckhout & Kircher (2017); Hagedorn et al. (2017).

<sup>50</sup>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 (10).

<sup>51</sup>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.

<sup>52</sup>Proof: Appendix A.

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  $t$ . The individual's expected change in wages can be summarized by:

$$\begin{aligned}
 E[y_{it} - y_{it-1} | j(i, t) = 2, j(i, t-1) = 1] &= \psi_2 - \psi_1 + E[\varepsilon_{it} | j(i, t) = 2] - E[\varepsilon_{it} | j(i, t-1) = 1] \\
 &= \psi_2 - \psi_1 + E[\lambda_{i2} - \lambda_{i1} | j(i, t) = 2, j(i, t-1) = 1] \\
 &+ E[\omega_{it} | j(i, t) = 2] - E[\omega_{it} | j(i, t-1) = 1] + E[v_{it} | j(i, t) = 2] - E[v_{it} | j(i, t-1) = 1]
 \end{aligned} \tag{11}$$

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. \(2013\)](#)<sup>53</sup>. 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<sup>54</sup> 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<sup>55</sup>. This conclusion goes in line with similar findings in the labor literature<sup>56</sup>. 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 asymmetric plot. Focusing on the switching trajectories between first and fourth quartiles<sup>57</sup>,

<sup>53</sup>The event study methodology is also used in [Card et al. \(2015\)](#); [Finkelstein et al. \(2016\)](#); [Best et al. \(2017\)](#).

<sup>54</sup>Define firm where employee works in  $t-1$  as the "origin" firm. Define firm where employee works in  $t$  as the "destination" firm.

<sup>55</sup>The ratio is calculated between the two expected wage differentials from  $t-1$  to  $t+1$ .

<sup>56</sup>[Card et al. \(2013, 2015\)](#).

<sup>57</sup>The same patterns can be found for other trajectories; see Appendix E.

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 (11). In that case, a positive (or negative) drift in employees' quality<sup>58</sup> 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  $E[v_{it}|j = 2] \neq E[v_{it}|j = 1]$  in equation (11) and would translate into an Ashenfelter's dip<sup>59</sup> 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's career provide a case for the use of the non-managerial labor market spell of the CEO 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 2.1 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 (7) 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<sup>60</sup>. I cluster standard errors at the employee-firm level, accounting for the two-way fixed effects nature of the regression in equation (7). I decompose the

<sup>58</sup>e.g. Increases in human capital accumulation.

<sup>59</sup>Ashenfelter (1978).

<sup>60</sup>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.

variance and covariance components log wages as:

$$\begin{aligned} Var(y_{it}) = & Var(\alpha_i) + Var(\psi_{j(i,t)}) + Var(X_{it}\beta) + 2 * Cov(\alpha_i, \psi_{j(i,t)}) \\ & + 2 * Cov(\psi_{j(i,t)}X_{it}\beta) + 2 * Cov(\alpha_i, X_{it}\beta) + Var(\varepsilon_{it}) \end{aligned} \quad (12)$$

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%<sup>61</sup>. 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<sup>62</sup>.

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<sup>63</sup> for firm  $j$  at time  $t$ :

$$q_{jt} = \delta + \phi CEO_{i(j,t)} + \mathbf{W}_{jt}\beta + \mathbf{Z}_{jt}\gamma + \epsilon_{jt} \quad (13)$$

where  $q_{jt}$  is the log of deflated sales,  $\mathbf{W}_{jt}$  is a vector of variable inputs and  $\mathbf{X}_{jt}$  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  $t$ ) also enters the production function as a state variable<sup>64</sup>, as suggested in section 2.1. The sequence  $A_{jt}$  is unobserved firm productivity. The error term  $\epsilon_{jt}$  has the following structure:

$$\epsilon_{jt} = A_{jt} + \eta_{jt} \quad (14)$$

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

<sup>61</sup>These results go in line with the labor economics literature: Card et al. (2013); Bonhomme et al. (2017b).

<sup>62</sup>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.

<sup>63</sup>Bloom & Van Reenen (2007).

<sup>64</sup>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<sup>65</sup> in their treatment of simultaneity bias<sup>66</sup>. Both works resort to proxy variables (investment and materials, respectively) to measure firm productivity in two step estimation approaches<sup>67</sup>.

Following recent literature, I use a GMM approach rather than two step procedure to jointly estimate both firm productivity and input coefficient<sup>68</sup> in equation (13). 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<sup>69</sup>. 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.  $\hat{\alpha}_i$ <sup>70</sup>. 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. I consider intermediate materials and services used as variable inputs<sup>71</sup>.

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 polynomial on both  $x_{jt-1}$  and  $m_{jt-1}$ . See Appendix C for further details regarding the GMM estimation.

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<sup>72</sup> overidentification test to assess the joint validity of

<sup>65</sup>Wooldridge (2009) methodology also takes into account the critiques in Blundell & Bond (2000) and Akerberg et al. (2006).

<sup>66</sup>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.

<sup>67</sup>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.

<sup>68</sup>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.

<sup>69</sup>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.

<sup>70</sup>The standardized person fixed effects are given by  $\frac{\hat{\alpha}_i - \mu_\alpha}{\sigma_\alpha}$  and  $\mu_\alpha = \frac{\sum_{CEO} \hat{\alpha}}{n_{CEO}}$ .

<sup>71</sup>In Appendix B, I further detail how each input is measured

<sup>72</sup>Sargan (1958) and Hansen (1982).

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 instruments 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, measured as Research and Development expenditures. 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<sup>73</sup>.

## 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.

---

<sup>73</sup>Pérez-González (2006); Bennedsen et al. (2007); Bandiera et al. (2013).

## 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 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 7 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 their individual 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 over-estimated<sup>74</sup> and, consequently, so will the coefficient of CEO quality in the second stage regression. Results can be found in Table 8. 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 11%, compared to 5% when using a proxy measure that has been significantly clean of endogeneity.

My findings go 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 8.

## 5.2 Production Functions

I estimate alternative specifications for the production function<sup>75</sup>. I use an OLS estimation of equation (13) 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

---

<sup>74</sup>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.

<sup>75</sup>[Petrin & Sivadasan \(2013\)](#).

different from 1:

$$q_{jt} = \sum_{k=1}^5 \beta_k X_{jt}^k + \beta_{kk} X_{jt}^{k^2} + \sum_{l \neq k} \sum_k \beta_{lk} X_{jt}^k X_{jt}^l + \epsilon_{jt} \quad (15)$$

where  $q_{jt}$  represents deflated log of sales for firm  $j$  in year  $t$ ,  $X_{jt}$  stands for one of the five input variables<sup>76</sup>. As in the OLS specification, I use the same with firm  $\times$  year fixed effects.

In Table 10, 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

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 key challenges in the estimation of a two-sided heterogeneity model in the presence of complementarities in the context of these data. First, CEO job mobility shows signs of endogeneity; that is, the CEO appears to change jobs according to, besides CEO and firm types, the CEO-firm match-specific wage realizations, either past or expected in the future. This directly affects the measuring of complementarities. Second, accounting for complementarities demands a more flexible and parsimonious model than the fixed effects 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 the two<sup>77</sup>. Both these challenges are addressed in the finite mixture model developed by Bonhomme et al. (2017b) (BLM for short). The novelty in this section is that I apply their model to a setting of revenue productivity estimation based on CEO-firm matches. The dynamic model I use includes a one-period Markov process for job mobility, revenue path dependence and allows for an unrestricted form of complementarity between the firm and CEO<sup>78</sup>. Their framework enhances parsimony in comparison to the fixed effects model, by consistently estimating a small number discrete heterogeneity classes through a dimension reduction technique via *kmeans* clustering.

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 parameters allow me to characterize firm productivity distribution by CEO, firm and

<sup>76</sup>Labor, capital, materials, services and CEO quality.

<sup>77</sup>Introducing a new parameter in a fixed effects regression is not an advisable option, since the incidental parameter bias discussed in section 4.4 would be further exacerbated in this setting given the (even more) limited amount of observations of each CEO-firm match.

<sup>78</sup>Rather than fixed parameters, probabilities are estimated.

match heterogeneity and also permits resorting to counterfactual experiments. I run such counterfactual experiment to establish the effect in firm productivity when match complementarities are artificially broken by randomly reassigning CEOs to firms.

## 6.1 Dynamic 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 the identifier  $m_{it}$ , which is equal to 1 if the CEO switches firms from time  $t$  to time  $t+1$  and 0 otherwise. 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  $k_{it} \in \{1, \dots, K\}$ . CEO heterogeneity is also discrete and finite:  $\alpha_i$  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  $t = (0, \dots, T)$  and a stream of inputs represented by a vector  $\mathbf{X}_{jt}$ .

I focus on a dynamic model<sup>79</sup> in which both job mobility and employee's earnings exhibit a specific type of serial correlation. The dynamic model has 4 periods. The CEO moves from firm  $k$  to  $k'$  between  $t=2$  and  $t=3$ . The 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  $\alpha$  and firm class  $k$  and  $k'$ , but also on match specific revenue realizations  $w_{i2}$ . However, it cannot depend on  $w_{i1}$ .
2. Revenues  $Y_{j,t+1}$  depend on the former period revenues realization,  $Y_{jt}$  but not on  $Y_{j,t-1}$ .

The assumption detailed above<sup>80</sup> 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:

$$y_{jt} = \rho_t y_{j,t-1} + a_{1t}(k_{i,t-1}) + a_{2t}(k_{it}) + b_t(k_{it})\alpha_i + \mathbf{X}_{jt}\mathbf{c}_t + \varepsilon_{jt} \quad (16)$$

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

## 6.2 Identification

Identification in this model relies, as before, on job mobility. BLM show that the key condition for identification of the model described in 6.1 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

<sup>79</sup>BLM discuss a static and a dynamic version of the model. Given that the static model is equivalent to the fixed effects approach used in section 4 in terms of its assumptions, I focus on BLM's dynamic model as it provides a relaxation of the assumptions of no endogeneity in mobility.

<sup>80</sup>A formal representation of these assumptions is included in Appendix D.

we are aiming to estimate CEO-firm complementarities. To do so, we need sufficient variation in the latent types of job movers between different firm classes<sup>81</sup>.

Identification follows the same steps as the estimation. First, a dimension reduction *kmeans* 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  $\alpha_i$  with  $i = \{1, \dots, L\}$  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. Grouping firm heterogeneity into clusters can be accomplished through a machine learning classification problem. BLM propose clustering the  $J$  firms in the sample into classes of log earnings distribution by solving the a weighted *kmeans*<sup>82</sup>. I adapt this problem to the setting of firm revenues:

$$\min_{k(1), \dots, k(J); H_1, \dots, H_k} \sum_{j=1}^J n_j \int (F_j(\hat{y}) - H_{k(j)}(y))^2 \quad (17)$$

where  $n_j$  is the number of CEO<sup>83</sup> 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. For the *kmeans* algorithm, I use all observations for each CEO-firm pair to establish the class of the firm, regardless of time period. The *kmeans* algorithm procedure is explained in Appendix D.

Table 11 presents the descriptive statistics at the firm class level. As in BLM, I use  $K=10$ , although in robustness checks I expand the number of clusters to 15 and 20. Notice that average salary and average gross revenue are increasing in firm class. Other

<sup>81</sup>In particular, complementarities would not be identified if there were random assignment of CEOs to firms.

<sup>82</sup>The *kmeans* algorithm belongs to a class of unsupervised learning algorithms. Unsupervised learning is indicated when the econometrician does not have prior knowledge on which classes to attribute each observation. Unsupervised learning 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 and, as BLM, assume this number is known. I use the Euclidean distance k-means algorithm as there is evidence that it performs best ([Singh et al., 2013](#)).

<sup>83</sup>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.

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 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:

$$\sum_{i=1}^{N^m} \sum_{k=1}^K \sum_{k'=1}^K \mathbf{1}\{\hat{k}_{i2} = k\} \mathbf{1}\{\hat{k}_{i3} = k'\} \times \ln\left(\sum_{\alpha=1}^L p_{kk'}(\alpha) f_{y_2, k\alpha}^f(Y_{i1}) f_{kk'\alpha}^m(Y_{i2}, Y_{i3}) f_{y_3, k'\alpha}^f(Y_{i4})\right) \quad (18)$$

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 (18) recovers log-revenues densities for each CEO type  $\alpha$ . 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 ( $t=2$ ):

$$\sum_{k=1}^K \sum_{k'=1}^K \mathbf{1}\{\hat{k}_{i2} = k\} \times \ln\left(\sum_{\alpha=1}^L q_k(\alpha) f_{y_2, k\alpha}^f(Y_{i1}) f_{kk'\alpha}^s(Y_{i2}, Y_{i3}) f_{y_3, k'\alpha}^f(Y_{i4})\right) \quad (19)$$

Both equations (18) and (19) are single-agent correlated random-effects log-likelihood functions. Though identification of  $k$  and  $\alpha$  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  $\alpha$  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<sup>84</sup>. Note that, if CEO and firm allocation is random, then the term  $b_t(k_{it})\alpha_i$  in equation (16) 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:

$$E[Y_i] - E^{cf}[Y_i] = E[b(k_i)\alpha_i] = cov(b(k_i), E[\alpha|k_i]) \quad (20)$$

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 (20) is positive and therefore there is a relationship between CEO type and complementarities that will not be negligible in the data simulations.

The expected change in average productivity is -2%; 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. \(2017a,b\)](#) framework to analyze CEO-firm complementarities in firm production<sup>85</sup>. I find the BLM model presents a very innovative approach for the estimation of two-sided heterogeneity models and 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 replicable in other matched panels, which are becoming increasingly available to econometricians.

<sup>84</sup>This counterfactual exercise abstracts from equilibrium conditions.

<sup>85</sup>Other important models, such as [Arellano & Bond \(1991\)](#) provide a generalization of the reduced-form approach to include path dependence in wages. The more recent [Hagedorn et al. \(2017\)](#) or [Abowd et al. \(2017\)](#) use structural (former) or bayesian (latter) approaches to explain CEO-firm match complementarities and sorting. Further discussion on the connection between reduced-form and structural approaches in the context of two-sided heterogeneity models can be found in [Appendix D](#).

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 has some limitations. While 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 imposing restrictions on heterogeneity. In the case of this model, random effects leaves way for unrestricted CEO-firm complementarities, but restricts heterogeneity to a small finite support. This may be an issue, the number of points in the support of both firm and CEO heterogeneity is a difficult issue that has received much attention 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. Rather, one can think that it is plausible that firm heterogeneity classes are more fluid and behave as probabilistic distributions over types, as assumed for the CEOs.

Overall, I believe this model has significant traction in the data setting at hand and points to plausible conclusions regarding CEO-firm complementarities in production, maintaining a fair amount of degrees of freedom in the model specification and parsimony. The two-step approach significantly reduces computation burden while providing consistent estimates for CEO and firm heterogeneity.

## 7 Conclusions

In this paper I present evidence that CEO heterogeneity type, or “quality”, is important both for the overall determination of variation in firm productivity across firms and for the within firm type productivity variation, given by the CEO-firm complementarity in production. 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, a counterfactual experiment estimates that complementarities explain between 2% 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. This means that improving establishment productivity must take into account the role of the CEO and the observable variables that are connected CEO quality, but also the role of CEO-firm complementarities.

Knowledge of the impact of CEO quality for firm performance has important policy implications. First, from an equity point of view, it contributes to shed more light

on the debate regarding the size and recent increase in wage inequality<sup>86</sup> and the size and increase in CEO pay<sup>87</sup>. It seems that, while there is ample evidence in the literature that the rise in firm size has contributed disproportionately for the rise in CEO wages, I find evidence that a non-negligible part firm revenue productivity can be attributed to the complementarities between the CEO and the firm. 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<sup>88</sup>.

---

<sup>86</sup>See Card et al. (2013); Song et al. (2015).

<sup>87</sup>See Murphy & Zabojnik (2004); Gabaix & Landier (2008); Terviö (2008); Frydman & Saks (2010).

<sup>88</sup>An example could be including, as part of the recruitment requirements, explicit soft skill analysis of complementarity between the managerial style and abilities and the firm managerial policies. In fact, one can observe this a tend in recent years, in particular for very large firms such as Google. The results in this paper suggest that smaller firms also stand to benefit from this practice.

## References

- Abowd, J. M., Creedy, 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).
- Akerberg, 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.
- Arellano, M., & Bond, S. (1991). Some tests of specification for Panel Data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277–297.
- 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., & Van Reenen, J. (2007). Measuring and Explaining Management Practices across Firms and Countries. *The Quarterly Journal of Economics*, 122(4), 1351–1408.
- Blundell, R., & Bond, S. (2000). Gmm Estimation with Persistent Panel Data: an Application to Production Functions. *Econometric Reviews*, 19(3), 321–340.
- 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.
- MacQueen, J., et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 281–297).
- 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.
- Singh, A., Yadav, A., & Rana, A. (2013). K-means with three different Distance Metrics. *International Journal of Computer Applications*, 67(10).
- 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.

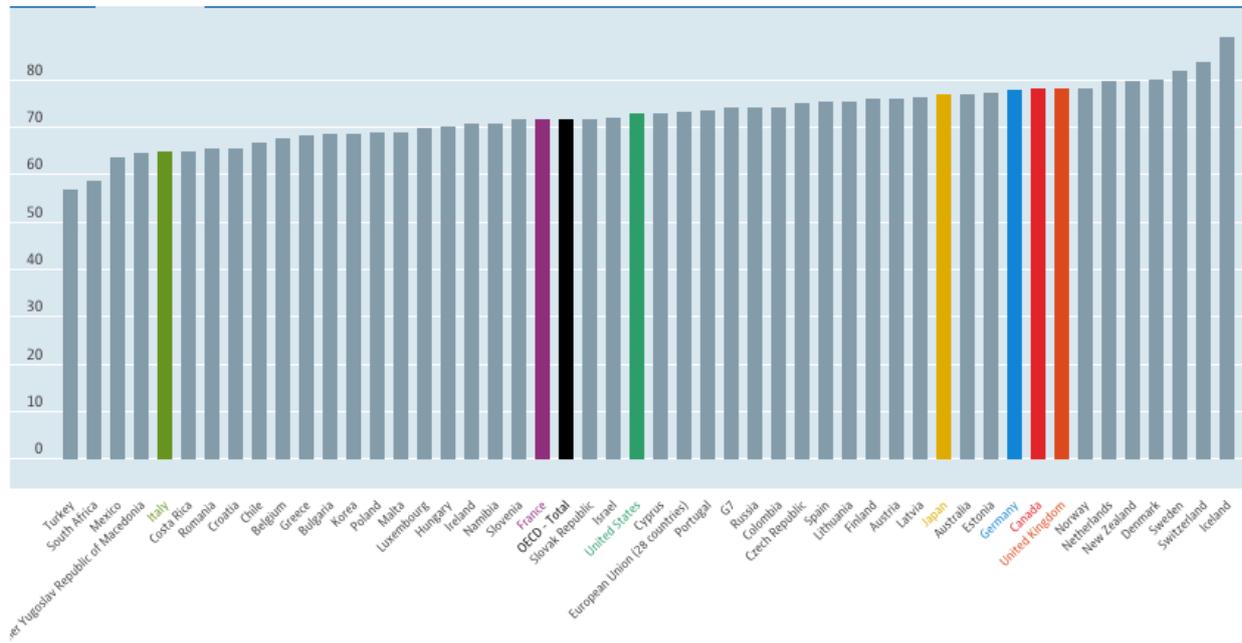
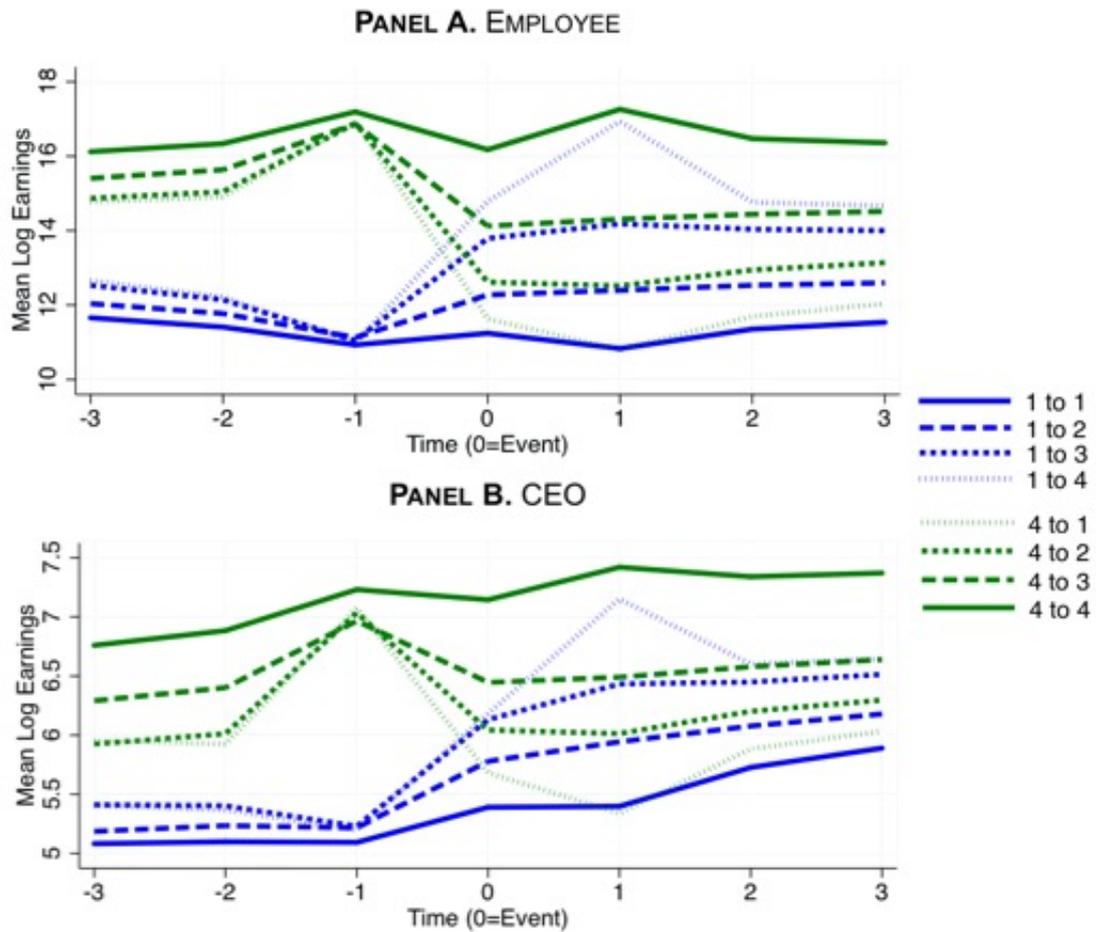
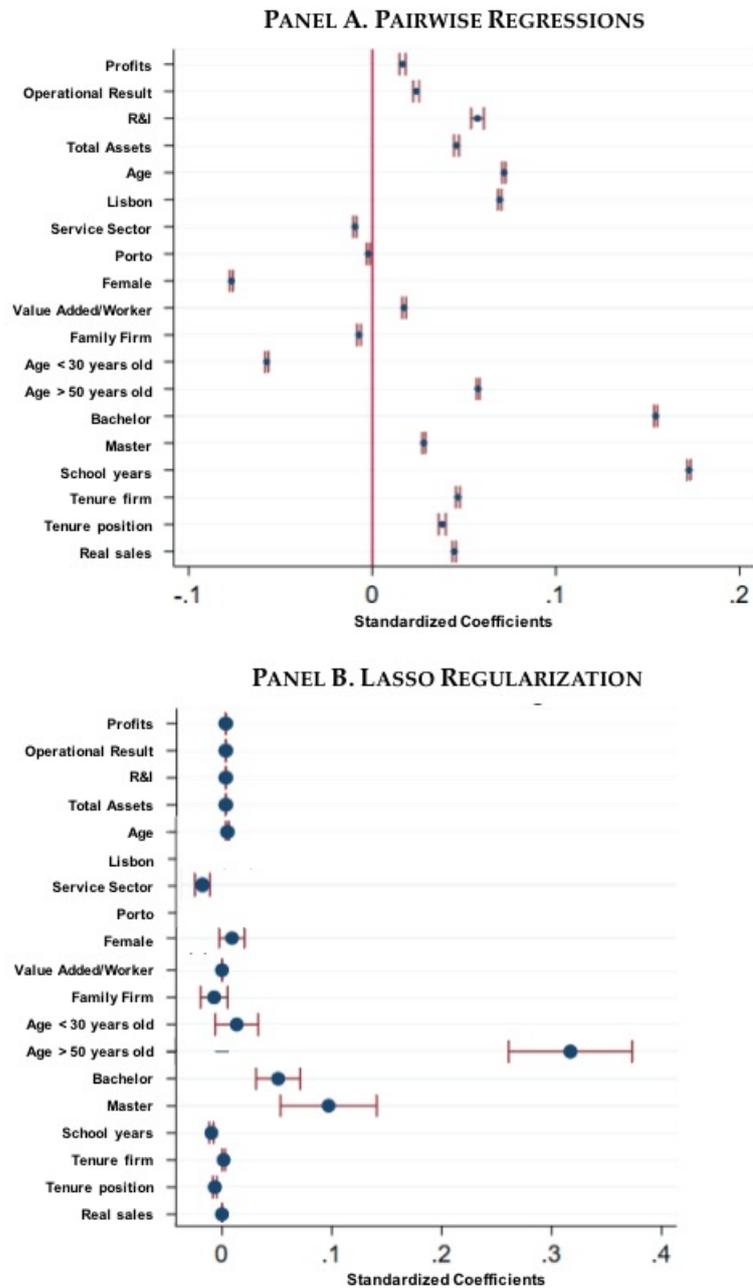


Figure 2: Event Studies - Employee and CEO Job Transitions.



<sup>88</sup>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, from 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.



<sup>88</sup>The figure exhibits two panels. Panel A presents the bivariate regression coefficient of estimated CEO fixed effects, estimated from the regression  $y_{it} = \alpha_i + \psi_{j(i,t)} + \mathbf{X}_{it}\beta + \varepsilon_{it}$ , 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		
	(1)	(2)	(3)	(4)	(5)	(6)
	Mean	St. Dev.	Obs.	Mean	St. Dev.	Obs.
<i>Demographics</i>						
Female (=1)	30.57%	0.46	2,976,326	34.44%	0.48	12,553,863
Age	42.50	11.68	2,976,326	36.78	12.24	12,553,863
Below 30 y.o. (=1)	18.23%	0.38	2,976,326	18.65%	0.39	12,553,863
Between 30 and 50 y.o.(=1)	50.99%	0.51	2,976,326	39.03%	0.49	12,553,863
Above 50 y.o. (=1)	30.78%	0.46	2,976,326	42.31%	0.49	12,553,863
<i>Education</i>						
Bachelors degree (=1)	29.78%	0.398	2,976,326	9.51%	0.27	12,553,863
Masters degree (=1)	3.04%	0.101	2,976,326	0.27%	0.05	12,553,863
<i>Tenure, Wages and Job Mobility</i>						
Tenure position	6.17	6.22	2,465,925	5.98	7.79	10,665,774
Log-wages	10.24	0.67	2,180,803	7.58	0.59	7,829,870
Job Mobility	2.15	1.58	2,976,326	3.45	1.49	12,553,863

Table 2: Descriptive Statistics - Largest Connected Set.

	CEO			Employees		
	(1)	(2)	(3)	(4)	(5)	(6)
	Mean	St. Dev.	Obs.	Mean	St. Dev.	Obs.
<i>Demographics</i>						
Female (=1)	25.57%	0.43	402,726	42.47%	0.49	12,302,786
Age	40.33	6.50	402,726	32.29	7.53	12,302,786
Below 30 y.o. (=1)	38.43%	0.38	402,726	23.41%	0.34	12,302,786
Between 30 and 50 y.o.(=1)	40.79%	0.51	402,726	56.60%	0.49	12,302,786
Above 50 y.o. (=1)	20.78%	0.46	402,726	20.20%	0.50	12,302,786
<i>Education</i>						
Bachelors degree (=1)	33.78%	0.398	402,726	10.01%	0.27	12,302,786
Masters degree (=1)	7.69%	0.101	402,726	0.51%	0.50	12,302,786
<i>Tenure, Wages and Job Mobility</i>						
Tenure position	5.18	4.55	246,407	5.06	4.81	10,402,459
Log-wages	11.56	0.67	270,884	8.01	0.59	7,673,272
Job Mobility	2.49	1.84	402,726	3.52	1.54	12,302,786

Table 3: 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 4: Descriptive Statistics - Firm.

	<b>Firm</b>		<b>Largest Connected Set</b>	
	(1) Mean	(2) St. Dev.	(3) Mean	(4) St. Dev.
<i>Demographics</i>				
Lisbon (=1)	28.90%	0.453	29.70%	0.457
Porto (=1)	21.96%	0.414	24.09%	0.428
Manufacturing (=1)	36.64%	0.482	44.15%	0.496
Construction Sector (=1)	14.38%	0.351	13.16%	0.338
Services (=1)	48.98%	0.500	42.68 %	0.495
<i>Financials</i>				
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 5: Variance Decomposition - Wage Variation on Worker and Firm Heterogeneity.

	1986-2013		2005-2013	
	(1) Var	(2) Share	(3) Var	(4) 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%

Table 6: Variance Decomposition - Finite Sample Bias Adjusted.

	1986-2013	2005-2013
	(1)	(2)
Log-wages	100.00%	100%
Employee FE	53.10%	51.19%
Firm FE	18.68%	18.91%

<sup>88</sup>Table 5 displays the results of a variance decomposition exercise conducted as per equation 12. 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.

<sup>88</sup>Table 6 displays the results of a variance decomposition exercise corrected with a bootstrap method, detailed in Appendix C.

Table 7: 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
<b>CEO Proxy</b>	<b>0.049</b>	0.01	-	<b>0.046</b>	0.00	-
Employees	0.14	0.01	-	0.13	0.00	-
Capital	0.05	0.00	-	0.06	0.01	-
Materials	0.35	0.00	-	0.37	0.02	-
Services	0.36	0.00	-	0.32	0.00	-
# Observations	1,220,771			1,660,193		

<sup>88</sup>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 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 8: Placebo Production Function Estimates - Random Employee.

	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
<b>Random Employee</b>	0.21	0.78	-	0.04	1.01	-
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	-

<sup>88</sup>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. Columns 3 and 6 present the Sargan-Hansen overidentification test p-values.

Table 9: 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
<b>CEO FE</b>	<b>0.09</b>	0.07	-	<b>0.10</b>	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	-

<sup>88</sup>Table 9 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 10: Alternative Production Function Specifications.

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

<sup>88</sup>Table 10 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.

Table 11: Descriptive Statistics - Firm Class Analysis Results.

Class	1	2	3	4	5	6	7	8	9	10	All
Avg. Age	30.97	43.88	44.20	44.25	44.21	42.79	44.06	43.68	43.30	43.10	<b>44.09</b>
Avg. Tenure	7.02	6.28	7.32	7.51	7.27	5.21	6.85	6.47	6.47	5.96	<b>7.13</b>
Avg. Salary (eur)	918.87	1,006.02	1,044.28	1,212.20	1,398.17	1,515.50	1,693.19	2,002.25	2,068.67	2,320.53	<b>1,365.27</b>
% BA	10.88%	14.47%	10.68%	12.19%	14.79%	22.14%	19.06%	20.99%	24.73%	27.70%	<b>24.65%</b>
% MA	0.71%	0.96%	0.78%	0.86%	1.00%	1.90%	1.34%	1.41%	1.92%	1.98%	<b>1.05%</b>
Avg. Revenues	376k	390k	546k	1,028k	1,187k	1,326k	2,301k	3,585k	5,406k	16,420k	<b>1,531k</b>
# Employees	84,213	28,897	156,017	199,917	182,429	4,575	128,530	68,344	28,670	6,524	<b>888,116</b>
% Manufacturing	20.87%	19.93%	23.16%	21.52%	17.09%	19.23%	15.26%	13.15%	30.12%	9.49%	<b>18.77%</b>
% Construction	20.23%	21.32%	20.47%	18.35 %	14.63%	23.04%	13.04%	14.78%	38.35%	22.07%	<b>17.23%</b>
% Services	58.89%	58.75%	56.37%	60.12%	68.72%	57.73%	71.69%	72.06%	44.90%	68.44%	<b>64.01%</b>

<sup>88</sup>Table 11 presents the descriptive statistics for each estimated firm class, according to the *kmeans* clustering algorithm presented in section 6.3.

## A Appendix: Proofs

### A.1 CEO Quality and Firm Productivity with Random Assignment

Proof of Proposition 1:

$$\frac{dPY_{j,t}}{d\alpha_i} = \frac{d(P * A_j \alpha_i^\mu L_{j,t}^\delta K_{j,t}^{1-\delta-\mu})}{d\alpha_i} = \mu > 0, QED \quad (21)$$

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

Proof of Proposition 2:

$$\frac{dPY_{j,t}}{d\theta_{ij}} = \frac{d((P * A_j) \theta_{ij} \alpha_i^\mu L_{j,t}^\delta K_{j,t}^{1-\delta-\mu})}{d\theta_{ij}} = (P * A_j) \alpha_i^\mu L_{j,t}^\delta K_{j,t}^{1-\delta-\mu} > 0, QED \quad (22)$$

where the last step derives from the fact that input levels are all positive and firm's productivity term,  $A_j$  must also be positive.

### A.3 Causal Identification

In section 4.3 I claim that consistent estimation of equation (9) via OLS implies the following assumptions regarding the interaction of the error terms with explanatory variables:

$$E[e^{i'} \varepsilon] = 0, \forall i \quad E[f^{j'} \varepsilon] = 0, \forall j \quad E[x^{k'} \varepsilon] = 0, \forall k \quad (23)$$

Furthermore, given the assumptions made on the error term components, causal identification of equation (9) boils down to the verification of the assumption  $E[f^{j'} \varepsilon] = 0, \forall j$ , because this equation is a direct result from the assumptions on the error terms.

A sufficient condition for this equation to hold is that the assignment of employees to firms is strictly exogenous with respect to  $\varepsilon$ :

$$P(J(i, t) = j | \varepsilon) = P(J(i, t) = j) = G_{jt}(\alpha_i; \psi_1, \dots, \psi_J) \forall i, t \quad (24)$$

where the employment probability functions  $G_{jt}$  sum up to 1 at every period for every employee  $i$ . This does not preclude sorting among  $\alpha$  and  $\psi$ .

## B Appendix: Data & Sample Selection

### B.1 CEO Quality

In the context of this paper, CEO quality or ability is defined as the unobserved heterogeneity component that influences the firm's production function. This unobserved

heterogeneity is considered a “black box” that has been studied in the management 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 relating it to unobserved heterogeneity, I do not aim to present a definition of manager quality from a psychological perspective of skill multiplicity; rather, I consider quality as an identity factor that the CEOs bring with them throughout their career, before and after becoming CEOs.

## B.2 Identifying General Managers or CEO

The definition of “manager” or managerial position within a firm, albeit seemingly intuitive in the business world, is not straightforward in a scientific context. An accurate definition of the manager is a key step in the analysis of their quality and/or impact. As such, careful consideration should be granted to pinning the precise notion of “manager” to be used. In this paper, I focus on the analysis of the impact of the so-called general manager, director, or CEO, in firm performance. For conciseness, I broadly define this highest level manager as “CEO” throughout the paper. I consider a practical definition postulated by the ILO<sup>89</sup>, whereby I consider the code 112 (Managing Directors and Chief Executives).

The data set *Quadros de Pessoal* contains a variable called *job title*<sup>90</sup> that corresponds to a 6-digit occupational classification system which was implemented as of 1995 (CNP 94). This occupation code divides top-managers into two classes: 12 corresponds to medium and large sized firm Directors and 13 corresponds to small firm Directors, where small firms are defined as having fewer than 10 workers. The class of Firm Directors (12) is further detailed into different categories: General Managers, Operations Managers, and Other Managers<sup>91</sup>. Moreover, these data contain another important variable for hierarchical classification, *qualification*<sup>92</sup>. This is a categorical variable which takes value=1 if the employee ID belongs to the highest hierarchical class within the firm. This variable is connected to the legal form of contract associated to the employee. Finally, the *Quadros de Pessoal* data set also contains a variable entitled *professional status*<sup>93</sup>, a categorical variable that takes value 1 if the observation (employee ID) is also the owner of the firm.

I identify the General Manager (or CEO) of each firm by resorting to a classification procedure<sup>94</sup> in which different criteria are used at each step:

**General Manager.** For each firm-year pair, I use the variable *job title* attribute the “CEO” label to the employee ID who is identified as General Manager. If there is a

---

<sup>89</sup>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>.

<sup>90</sup>The variable *job title* is called “Profissão” in the *Quadros de Pessoal* data set.

<sup>91</sup>Administrative, Financial and Sales Managers.

<sup>92</sup>The variable *qualification* is called “qualif” in the *Quadros de Pessoal* data set.

<sup>93</sup>The variable *professional status* is called “sitpro” in the *Quadros de Pessoal* data set.

<sup>94</sup>This classification procedure as well as sample selection is inspired by those used in Queiró (2016).

tie<sup>95</sup>, I pick the manager with the highest salary among the ones classified as General Manager. This step identifies a total of 12.83% of the firm-year observations.

**Operational Manager.** Using the same variable, *job title*, to identify the Operational Manager in the absence of an identified General Manager. I attribute the “CEO” label to the identified Operational Manager in this case. If there is a tie<sup>96</sup>, I pick the manager with the highest salary among the ones classified as Operational Manager. With this additional step, I can identify a total of 30.37% of the firm-year observations.

**Other Manager.** Using the same variable, *job title*, to identify Other Managers in the absence of an identified General or Operational Manager. I attribute the “CEO” label to the identified Other Manager in this case. If there is a tie<sup>97</sup>, I pick the manager with the highest salary among the ones classified as Other Manager. With this additional step, I can identify a total of 41.49% of the firm-year observations.

**Owner.** After fully exploiting the *job title* variable, I turn to ownership status to identify the head of the firm for the remaining firm-year pairs with unidentified General Manager or CEO. I use the variable *professional status* described above and label as “CEO” the employee ID for whom this variable takes value 1 (corresponding to the owner). There are no unsolved ties using this criterion. With this additional step, I can identify a total of 51.65% of the firm-year observations.

**Top Hierarchical Class.** I now use the variable *qualification* described above to identify remaining firm-year pairs. I attribute the label “CEO” to the employee ID associated with the highest hierarchical class within *qualification* (*qualification*=1). If there is a tie<sup>98</sup>, I pick the manager with the highest salary among the ones classified as Operational Manager. With this additional step, I can identify a total of 61.21% of the firm-year observations.

**Previous Manager.** For remaining unidentified firm-year pairs, I attribute the label CEO to an employee ID which was classified as a General Manager/CEO (according to any of the criteria above) in the period before and is still employed at the same firm<sup>99</sup>. There are no unsolved ties using this criterion. With this additional step, I can identify a total of 67.17% of the firm-year observations.

**Next Manager.** Similarly, for remaining unidentified firm-year pairs, I attribute the label CEO to an employee ID which was classified as a General Manager/CEO (according to any of the criteria above) in the period after in the same firm<sup>100</sup>. There are no unsolved ties using this criterion. With this additional step, I can identify a total of 71.26% of the firm-year observations.

**Maximum Salary.** Finally, for any remaining firm-year pair with unidentified CEO, I pick the employee which displays the highest salary. There are no unsolved ties using this criterion. With this additional step, I can identify a total of 93.09% of the firm-year observations.

The remaining firm-year pairs with no identified General Manager/CEO pertain

---

<sup>95</sup>Ties happen for 0.54% of the data.

<sup>96</sup>Ties happen for 0.59% of the data.

<sup>97</sup>Ties happen for 0.06% of the data.

<sup>98</sup>Ties happen for 3.76% of the data.

<sup>99</sup>This indicates some type of error/gap in the survey completion.

<sup>100</sup>This indicates some type of error/gap in the survey completion.

to those that do not disclose personnel salaries.

### B.3 Sample Selection

*Quadros de Pessoal* is a mandatory annual survey that contains personnel information on any private sector firm that employs at least one individual by October of each year. It is an anonymized database with identifiers for both firm and employee and spans from 1982 to 2013, with two gaps (1999 and 2001), totalling 29 years. With an initial data set composed of a total 6,140,063 firm-year pairs and 62,661,660 employee level observations, after eliminating missing identifiers, I proceed to restrict the selection of the final sample to be used.

First, I identify public enterprises or partially state-owned organizations, according to two criteria. I label as *public firms* those whose percentage of public capital exceeds 50%. I attribute that classification to firms that are identified as *public administration* in *legal status* variable of this data set. I discard firms who are labelled as *public* at any point in time since hierarchical structures in the public sector are very different from the private sector, with little cross-sector or within-firm mobility.

Second, and based on the step 1, I identify firms who have at least 50 of the same employees as a firm formerly identified as public and that no longer appears in the data. This amounts to identifying privatized firms which often times maintain their public-style hierarchical structures.

Third, I winzorize the data on both salaries and firm revenues at the 99<sup>th</sup> percentile. I do not use firms with less than 20 employees.

Lastly, I eliminate two classes of sectors, in line with other matched employer-employee studies in the literature: Agriculture and Fishery, and the Banking Sector. I therefore focus on the sectors pertaining to Manufacturing Industry, Construction and Services.

The *Informação Empresarial Simplificada* (IES) is also a mandatory annual survey that reports on firm-level Balance Sheet and Profit & Loss statements. It includes information on firm's assets and liabilities, inputs, revenues, operating and financial profits, value added, sector, size (number of employees) and location. The survey is anonymized and contains a firm identifier that allows for the matching with the *Quadros de Pessoa* data set.

### B.4 Variables and Coding

In section 4.1 I use education fully interacted with a second-degree polynomial on age as time-varying covariates in the two-way fixed-effects model. **Education** is the number of completed schooling years and **age** calculated using the birthdates available in the data. In cases where different education attainment is recorded for the same employee, I take the mode of the education years reported for each employee. If there is more than one mode, I take the lowest. The outcome variable in this regression is monthly **salary**, calculated as a sum of base salary, bonuses and pay for extra hours.

In section 4.5 I use four input variables, taken from the IES data, besides the proxy for CEO quality calculated in section 4.1, to explain firm productivity. As variable in-

puts, I use **services** and intermediate **materials**. Both variables are taken directly from the firm's Profit & Loss Statement and deflated to 2000 euros according to two-digit firm sector price indices. **Labor** and **capital** are the state variables used in this section. Labor is measured as the number of non-CEO workers employed by the firm. I measure capital as deflated book value of fixed assets.

In section 4.5 I focus on a well-defined output that is transversally applicable across different economic sectors: deflated revenue productivity. Revenues from production or service delivery within the private sector are a primary evidence of firm performance. In fact, revenue productivity is frequently used in the Organization and Growth literatures (Bender et al., 2016; Hsieh & Klenow, 2009, 2014). I deflate revenues according to two-digit firm sector price indices.

## C Appendix: Econometric Model

### C.1 Largest Connected Set

In section 4.2, when constraining the data to the largest connect set of employee-firm pairs, my samples are slightly modified in two predictable ways. One is the fact that larger-sized firms become more represented in the largest connected set sub-sample, both measured in operating revenue and number of employees. The other is the fact that these firms seem to last slightly longer within the whole sample. Yet, these differences do not present a worrisome outlook on the representativeness of the fixed-effect analysis. The two-way fixed-effects approach relies on manager job mobility as a source of variation. Therefore, the most important attribute of the largest connected set is that the composition of manager characteristics does not change substantially from that of the whole sample. This seems to be the case when comparing the descriptive statistics of the two samples, detailed in Table 1 and Table 2.

### C.2 Intuition Behind Endogenous Mobility: An Example

To provide intuition for the consistent estimation of individual fixed effects referred to in section 4.3, let us consider the simple case in which there are two CEO and two firms, producing in two time periods,  $t = \{1, 2\}$ . Consider Figure 4. At  $t = 1$ , the wage functions based on equation (7) can be evaluated as follows

$$\begin{aligned} y_{i=1,t=1} &= \alpha_1 + \psi_1 + \mathbf{X}_{1,t=1}\boldsymbol{\beta} + \varepsilon_{1,t=1} \\ y_{i=2,t=1} &= \alpha_2 + \psi_2 + \mathbf{X}_{2,t=1}\boldsymbol{\beta} + \varepsilon_{2,t=1} \end{aligned} \tag{25}$$

where  $j(1, 1) = 1$  and  $j(2, 1) = 2$ . At  $t = 2$ , the wage functions based on equation (9) can be evaluated as follows

$$\begin{aligned} y_{i=1,t=2} &= \alpha_1 + \psi_2 + \mathbf{X}_{1,t=2}\boldsymbol{\beta} + \varepsilon_{1,t=2} \\ y_{i=2,t=2} &= \alpha_2 + \psi_1 + \mathbf{X}_{2,t=2}\boldsymbol{\beta} + \varepsilon_{2,t=2} \end{aligned} \tag{26}$$

where  $j(1, 2) = 2$  and  $j(2, 2) = 1$ . The expected change in output for employee  $i = \{1, 2\}$  from time period 1 to 2 moving from firm  $j$  to  $j'$  is given by a difference-in-difference analysis,

$$E[y_{i,t=2} - y_{i,t=1} | j_{i,t=2} = j', j_{i,t=1} = j] = \psi_{j'} - \psi_j + E[\varepsilon_{i,t=2} | j_{i,t=2} = j'] - E[\varepsilon_{i,t=1} | j_{i,t=1} = j] \quad (27)$$

Taking equation (8) into account, if  $\lambda_{i,j(i,t)}$ ,  $\omega_{it}$  and  $v_{it}$  are uncorrelated with job mobility from firm 1 to 2 or vice-versa (which is the determinant of the personal fixed effect), then  $E[\varepsilon_{i,t=2} | j_{i,t=2} = j'] - E[\varepsilon_{i,t=1} | j_{i,t=1} = j]$  can be cancelled out and the whole expression simplifies to the difference in firm fixed effects<sup>101</sup>,  $\psi_{j'} - \psi_j$ . Notice that, in case this holds, we should observe in the data that movements from firm  $j$  to  $j'$  should yield employees a symmetric gain as compared to movements from  $j'$  to  $j$ . This is what I test for in Figure 1.

### C.3 Finite Sample Issues: Variance Shrinkage

Estimating individual fixed effects in panel data sets entail an inherent challenge related to the panel structure. Each individual is observed a limited amount of instances in the time-series and therefore, the finite sample available for each individual may result in a generally upward incidental parameter bias<sup>102</sup> on estimated fixed effects.

Several corrections for this type of bias have been proposed in the literature. These corrections are often referred to as variance shrinkage procedures. In this paper, I use a method proposed by Kane & Staiger (2008) and used in Best et al. (2017). This method estimates sampling error via a bootstrap method and shrinks the variance of the fixed effects by that factor.

By this method, it is assumed that the estimated sample variance of employee and firm contains its true value ( $\sigma_\alpha^2$  and  $\sigma_A^2$ ) and a term that represents noise due to sampling error ( $\sigma_\xi^2$  and  $\sigma_\nu^2$ ) which arises from the finite nature of the individual sample:

$$Var(\hat{\alpha}) = \sigma_\alpha^2 + \sigma_\xi^2 \quad Var(\hat{A}) = \sigma_A^2 + \sigma_\nu^2 \quad (28)$$

where the parameters of interest are  $\sigma_\alpha^2$  and  $\sigma_A^2$ . I use a bootstrap technique to calculate standard errors of each of the two estimated fixed effects, which yields estimates of the sampling errors of this finite sample for both employee and firm fixed effects. Denote them  $s_e^2$  and  $s_f^2$ . I take the expected value of these estimated sampling error terms across employees,  $E_e(s_e^2)$ , and firms  $E_f(s_f^2)$ . This finally yields the following estimated variances:

$$\sigma_\alpha^2 = Var(\hat{\alpha}) - E_e(s_e^2) \quad \sigma_A^2 = Var(\hat{A}) - E_f(s_f^2) \quad (29)$$

The results of this variance shrinkage can be found in Table 5.

<sup>101</sup>In this context, firm fixed effects can be interpreted as firm-specific payment policies or incentives.

<sup>102</sup>See Andrews et al. (2008, 2012) for detailed explanations of this type of bias.

## C.4 Additively Separable: Sorting Tests

I test for the assumption of additive separability (log-linearity) assumed in the model described in section 4.1 by resorting to heat map tests. I take the firm and employee fixed effect, estimated in non-CEO years, and divide each of them into vintiles of the distribution. I then group each pair of vintiles corresponding to each observation and compute the average regression residual. I then plot the result into a color coded map.

Results can be found in Figures 5 and 6. No distinct pattern can be identified in either of the figures, which means that, after controlling for separately additive firm and employee heterogeneity, the remaining unexplained term is not driven by non-separable interactions between the two fixed effects.

## C.5 GMM Estimation of Firm Production Function

Since Marschak & Andrews (1944) first introduced the concept of simultaneity bias in the estimation of production functions, several attempts at correcting this problem have been suggested in the literature, the most widely used being Olley & Pakes (1996) (OP for short) and Levinsohn & Petrin (2003) (LP for short). OP show that investment can be used as a proxy variable for unobserved productivity by employing a two-step estimation method. They assume that the industry produces with a Cobb-Douglas technology and that factors underlying profitability differences among firms are neutral efficiency differences. The first step is the estimation of a production function that is linear in labour and non-parametric in  $g(A_{jt}, k_{jt})$  a function of productivity and capital, which are considered state variables. In a second step, to identify the capital elasticities, further assumptions need to be made. Here, the authors use Markov process assumptions on  $A_{jt}$ . They regress output net of labor (variable input) on capital and a consistent estimate of  $E[A_{jt}]$ .

LP propose modifications to the OP approach to address the problem of lumpy investment. The authors claim that evidence of costly adjustment to capital investments explains the fact that many firms present zero investment at certain years. This leads to kinks in the firm investment demand function, meaning that firms or plants may not swiftly respond to certain productivity shocks. In that case, correlation between the regressors and the error term can remain. If it is less costly to adjust the intermediate input (materials) and LP argue that it may respond more readily to productivity. They use intermediate inputs to proxy for unobserved productivity instead, maintaining the two-step estimation method.

More recently, Wooldridge (2009) uses both of the former works' proxy variables while implementing a Generalized Method of Moments (GMM) approach, which bypasses the two-step method and its subsequent need to bootstrap standard errors. I follow this approach as it is now standard in the literature. The estimated production function takes the form:

$$q_{jt} = \beta_{CEO}CEO_{jt} + \beta_l l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + \beta_s s_{jt} + \epsilon_{jt} \quad (30)$$

where  $A_{jt} + \eta_{jt}$  and  $A_{jt}$  is the persistent productivity term and  $\eta_{jt}$  is assumed to be an iid transitory shock to productivity. Results of this part of the estimation can be found

in section 4.5 and Table 6. The estimation of  $A_{jt}$  is non-parametric and instrumented by a function on lagged state variables and instruments  $g(X_{j,t-1}, m_{j,t-1})$ . The non-parametric estimation is conducted by approximating the function  $g(\cdot, \cdot)$  with third-degree polynomials in both state variables ( $CEO_{jt}$ ,  $l_{jt}$  and  $k_{jt}$ ) and instruments ( $m_{jt}$  and  $s_{jt}$ ).

## D Appendix: Finite Mixture Model

### D.1 Assumptions

Formal [Bonhomme et al. \(2017b\)](#) dynamic model assumptions, applied to the context of CEO-firm revenue distribution:

**Assumptions** (*dynamic model*)

1. *Job mobility*:  $m_{it}$ ,  $k_{i,t+1}$  and  $X_{j,t+1}$  are independent of  $Y_j^{t-1}$ ,  $m_i^{t-1}$  and  $X_j^{t-1}$  conditional on  $Y_{jt}$ ,  $\alpha_i$  and  $X_{jt}$ .
2. *Serial dependence*:  $Y_{j,t+1}$  is independent of  $Y_j^{t-1}$ ,  $k_i^{t-1}$ ,  $m_i^{t-1}$  and  $X_j^t$  conditional on  $Y_{jt}$ ,  $\alpha_i$ ,  $m_{it}$ ,  $k_{it}$ ,  $k_{i,t+1}$  and  $X_{j,t+1}$ .

[Bonhomme et al. \(2017b\)](#) discuss this and mobility assumptions at length in their paper.

### D.2 Reduced Form and Structural Models

The reduced form model of wage determination with two-sided heterogeneity put forth by [Abowd et al. \(1999\)](#) and used in section 4.1 of this paper relies on important and possibly limited assumptions. By relying on fixed effects to represent individual heterogeneity and therefore leaving it unrestricted, it strongly conditions how this heterogeneity enters the model of wage setting. The model does not allow for complementarities in wages between employee and firm, meaning that on average each employee should behave the same way at each firm in terms of his individual role in wage determination. [Abowd et al. \(1999\)](#) is also a static model, relying on the assumption that employee mobility is random after accounting for employee and current (only) firm type. On the other hand, structural models ([Shimer & Smith, 2000](#); [Postel-Vinay & Robin, 2002](#)) improve upon the strict [Abowd et al. \(1999\)](#) model assumptions by accounting for wage dynamics in mobility and match outcomes, but face significant computation challenges as they often rely on the estimation of a very large number of parameters to accommodate the flexibility in model assumptions.

[Bonhomme et al. \(2017b\)](#) present an innovative “hybrid” model which keeps the flavour of a fixed-effects model when accounting for firm-level heterogeneity -by grouping firms into discrete classes- and leaves the employee (CEO, in the case of this paper) type as random effect represented by a finite mixture model, therefore restricting heterogeneity at the employee level but leaving the CEO-firm complementarities unrestricted<sup>103</sup>.

---

<sup>103</sup>Note that the firm class fixed effect can also be viewed as a different type of random effect, since

### D.3 Kmeans Clustering

The *kmeans* (MacQueen et al., 1967) clustering is a type of unsupervised learning, which is used when you have unlabeled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable  $K$ . The algorithm works iteratively to assign each data point to one of  $K$  groups based on the features that are provided. Data points are clustered based on feature similarity. The *kmeans* algorithm is one of the simplest unsupervised learning algorithms that solve the clustering problem.

The procedure follows a simple method of classification of a given data set through a certain number of clusters (assuming the number of  $K$  clusters is known) fixed *a priori*. The main idea is to define  $k$  centers, one for each cluster. The algorithm places initial cluster centers as far away from each other as possible. The next step is to take each point belonging to a given data set that we want to classify and associate it to the nearest center, by computing the Euclidean distance<sup>104</sup>. When no data point is unclassified, the first step is completed. At this point we need to re-calculate  $k$  new centroids as barycenter of the clusters classified in the previous step. Step 1 is now repeated to refine the classification. These two steps result in a loop that is finalized when the average distance between each observation and its centroid is minimized.

The *kmeans* clustering mechanism contains a challenge from which I abstract in this paper. It requires the correct knowledge of the number of underlying discrete heterogeneity classes. There is a large literature dedicated to answering this challenge. In this paper, I assume prior knowledge in the number of classes and fix it at  $K = 10$  for the analysis presented.

### D.4 Finite Mixture Model

The finite mixture model provides a natural representation of heterogeneity in a finite number of latent classes<sup>105</sup> and translates into modelling a statistical distribution of a certain data sample as a weighted sum of different distributions. The idea behind this type of model, when applied to a matched CEO-firm dataset, is that the distribution of revenues is CEO-firm match type specific; that is, each distribution depends on the latent type of the CEO and the firm.

Experience suggests that usually only few latent classes are needed to approximate density well (Heckman). I assume  $L = 5$ , although I experiment with 6 and 7 (without significant changes to the results).

---

<sup>104</sup>The Euclidean distance is the most commonly used metric. Singh et al. (2013) discuss different distance metrics.

<sup>105</sup>Finite mixture models are also referred to as latent class models or unsupervised learning models.

## E Appendix: Additional Figures & Tables

Figure 4: CEO Mobility: Example in a Two Period, Two Firm Environment.

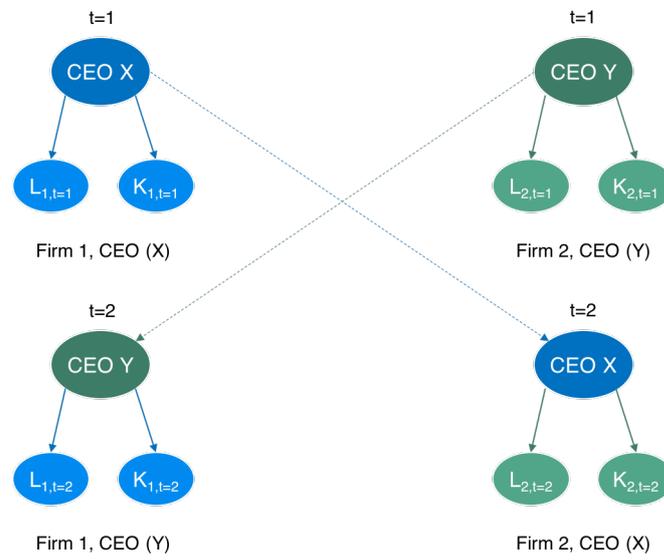


Figure 5: Heat Maps. Vingtiles of CEO and Firm Fixed Effects.

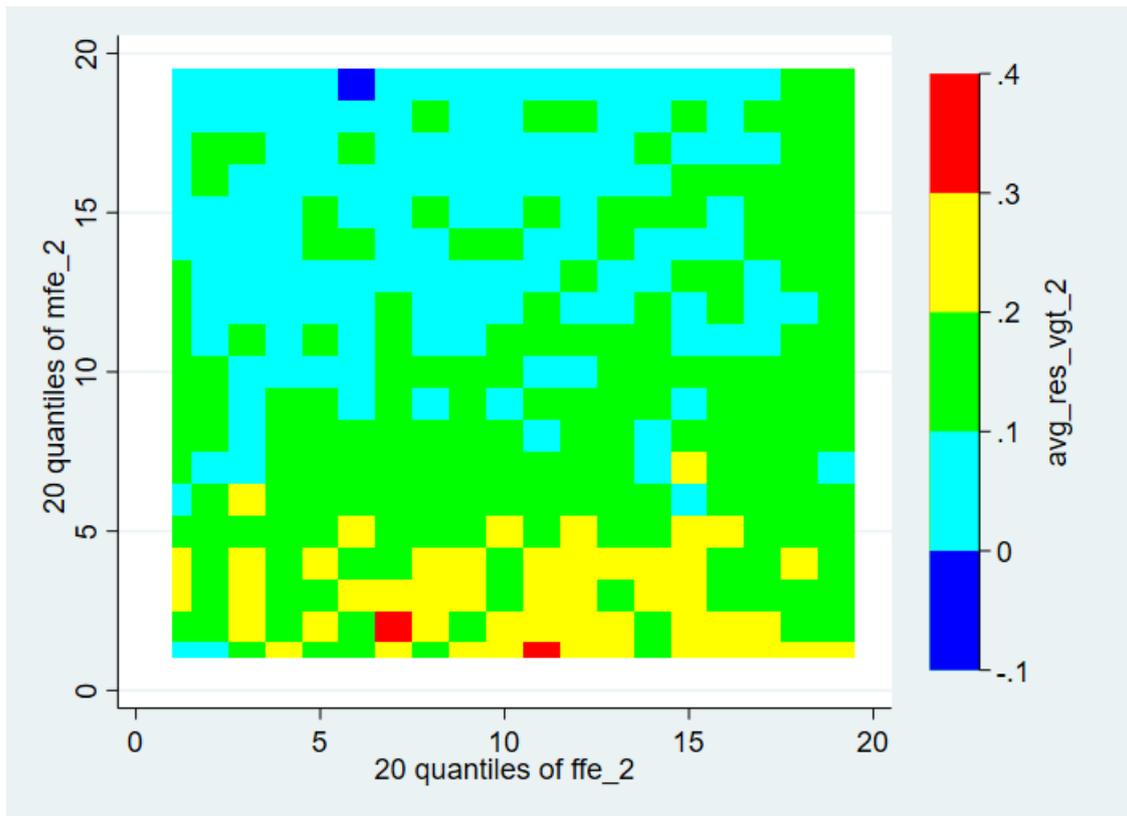


Figure 6: Heat Maps. Vingtiles of CEO and Firm Fixed Effects.

