CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.

BARGAINING POWER DYNAMICS: FAT TAIL DISTRIBUTION ANALYSIS

TESINA

QUE PARA OBTENER EL GRADO DE

MAESTRO EN ECONOMÍA

PRESENTA

GENARO BASULTO MEJÍA

DIRECTORA DE LA TESINA: DRA. SONIA BEATRIZ DI GIANNATALE MEGALLI

CIUDAD DE MÉXICO 2022

Este trabajo es para todos ustedes:
Mi mamá, María de la Luz;
Su esposo, Luis Leal;
Mi compañera de vida, Alexa Ximena;
Y para toda mi familia y las personas que me han apoyado para llegar hasta aquí.

Agradecimientos

Quiero agradecer a:

Mi asesora, la profesora Sonia Di Giannatale, por todo el apoyo que me ha dado a lo largo de la maestría;
Mi lector, Itza Curiel, por sus valiosos comentarios en mi trabajo;
Nuestro director del seminario de titulación, Irving Rojas, por todo el trabajo para que nuestro examen profesional fuera exitoso; Los profesores del CIDE, especialmente a Kaniska y Raciel, quienes me brindaron más que los conocimientos necesarios para obtener el grado;
A mis compañeros, especialmente a Daniela, por todos los trabajos y tareas en las que nos apoyamos, hiciste que el recorrido por el CIDE fuera más ligero;
A mi mamá, Luz, y su esposo, Luis, por apoyarme siempre en mi trayectoria y por darme y hacerme lo que soy hoy;
A Alexa Ximena, por siempre apoyarme y estar conmigo. Todo mi trabajo y mi exito es para ti.

Abstract

Drawing from a dynamic principal-agent model characterized by a state variable given by the agent's initial bargaining power and by a law of motion that governs the changes of their bargaining power, we introduce a fat tail distribution for the productivity possibilities in order to study heterogeneity in CEO compensations by analysing firm and industry-level characteristics. Through a numerical exercise, we find that the productivity possibilities determine CEOs' compensation size and distribution, but once this is determined, the bargaining dynamics within the firm shape the evolution of this compensation. We also present evidence that our model can describe to a good extent compensation behaviours observed in data.

Resumen

Partiendo de un modelo dinámico de agente-principal caracterizado por una variable de estado dada por el poder de negociación inicial del agente y por una ley de movimiento que gobierna los cambios de su poder de negociación, introducimos una distribución de cola ancha para las posibilidades de productividad con el fin de estudiar la heterogeneidad en las compensaciones de los directores generales mediante el análisis de características a nivel de la empresa y la industria. A través de un ejercicio numérico, encontramos que las posibilidades de productividad determinan el tamaño y la distribución de la compensación de los directores generales, pero una vez que esto está determinado, las dinámicas de negociación dentro de la empresa determinan la evolución de esta compensación. También presentamos evidencia de que nuestro modelo puede describir en buena medida los comportamientos de compensación observados en los datos.

Contents

List of Figures

List of Tables

1 Introduction

The size of the payments received by CEOs has been observed to be diverse amongst firms and to change through time (Frydman and Jenter, 2010). Explanations about the heterogeneity in CEO compensations amongst companies have relied on studying firm-level effects. Sur et al. (2015), for instance, present a model that predicts that the cash salary for CEOs is mostly driven by firm-specific factors, while Bouteska and Mefteh-Wali (2021) find that poor governance conditions lead to high compensation levels offered to CEO in US companies trough time. On the other hand, some studies have attributed this heterogeneity to industry-level factors, most commonly firm size (Gabaix and Landier, 2008, Edmans and Gabaix, 2009). However there is mixed evidence as to what is the real role that firm and industry-level factors play on determining CEO salary size and evolution. In this thesis, we aim to unite the two arguments by modeling the relationship between the company owners and its CEO accounting for bargaining power dynamics within the company, and a fat-tail productivity distribution that captures the heterogeneity between firm sizes.

Table 1: Compensation Summary.

	Variable	Mean	St. Dev.
1	EBIT	1045.47	3475.52
2	comp	41.12	176.62

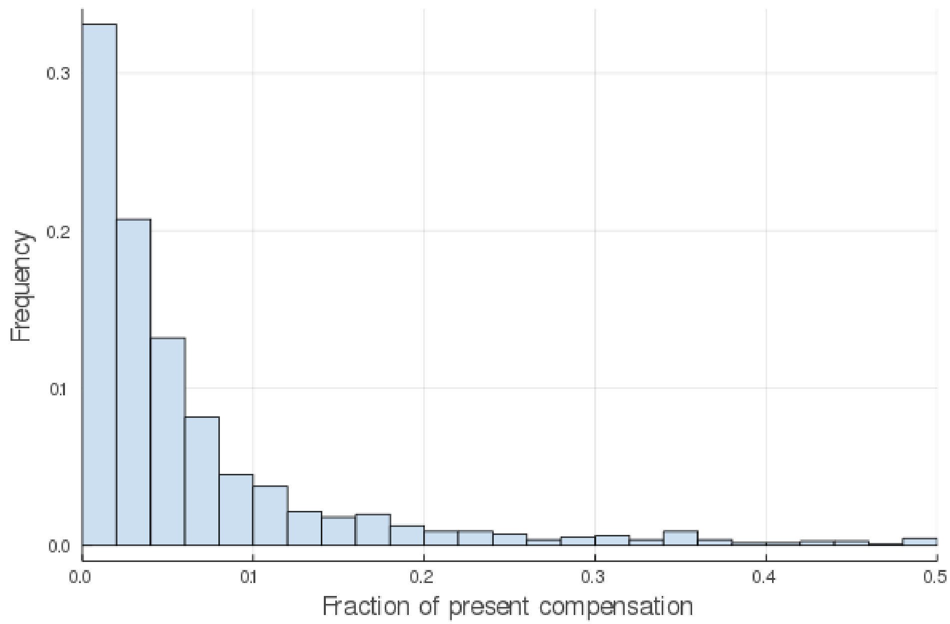
Source: Own Elaboration with data from S&P.

Table 1 presents a brief summary of data from the S&P Global Market Intelligence Trial at WRDS,[1] which contains annual financial details from around 300 companies in the US ranging from the year 1968 to 2014. The sample mean and standard deviation of the Earnings Before Interests and Taxes (EBIT) and the present compensation of the CEO (comp) amongst these companies are presented in this table. Notice that the average CEO earns \$41.12 MM annually, while the average EBIT amongst companies is of around \$1045.47. So, in average the CEO takes around 4% of the company's earnings as pay. However, if we look at Figure 1 we can see that a significant percentage of CEOs make more than 10% of the companies' EBIT with some

[1] We thank the people from WRDS at UPenn for letting us have access to this database. *This database is not available to the public.*

taking as much as 40%. This underlines the fact that there exists great heterogeneity among CEO compensation packages in the form of a number CEOs receiving a significantly higher pay, which could indicate higher ability to extract rent from shareholders (Murphy and Sandino, 2010) which implies a higher bargaining power for the CEO.

Figure 1: CEO compensation as a fraction of the company's EBIT.



Source: Own Elaboration with data from S&P.

To model the bargaining power dynamics between the companies' owners and its CEOs, we review a dynamic principal-agent model first presented in Di Giannatale et al. (2021) and solve it numerically using the Pareto family of probability distributions as the output distribution. In this model, the agent (CEO) bargaining power is the state variable which constitutes a departure from standard repeated principal-agent models, such as the one of Spear and Srivastava (1987). Given that agent compensation depends directly of the firm's output, using the Pareto distributions allows us to account for the *fat tail* behaviour in CEO compensation distribution observed in the real world. Our results indicate that the output distribution plays an important role on determining the level of risk and the size of the compensation that the CEO receives in the company. But once this size is determined, the bargaining power dynamics are the only factor that influences the evolution of the relationship.

We then validate our results using the aforementioned salary data from S&P where it is proved that industry-level factors such as firm size or industry sector do determine in some degree CEO salary. But heterogeneous salary trends arise within companies even with very similar characteristics, implying that salary trends are determined only by firm-level factors such as governance conditions. We also find that most CEOs in our sample take on a greater level of risk in the contractual relationship, which is something that our model captures and contrast with the lower levels of risk sharing predicted by classical principal-agent models.

The remaining of this work is organized as follows: In section 2, the related literature is reviewed to further motivate the topic. In section 3 we present the model. Section 4 details the numerical approach to the problem; first reviewing the Pareto family of probability distributions and underline the parameters used in the numerical solution and then presenting the results. In section 5 we use the S&P data and perform time series analysis on CEO salaries to later compare them with simulated salaries using our model. Finally, section 6 concludes the thesis.

2 Related Literature

The current work is related to the literature of dynamic contract theory and bargaining power in principal-agent relationships. Specifically, we aim to expand on the ideas of Di Giannatale et al. (2021) about the dynamics of the bargaining power in a long-term contractual relationship between a firm shareholders and its CEO. The main contribution of this thesis is to analyze the heterogeneity in CEO payments and their evolution among firms in the context of this model.

In the late 1970s, Hölmstrom (1979) presented a static model where a risk neutral principal and a risk adverse agent have a contractual relation subject to moral hazard. His model finds that when the agent's effort is not observable, optimal contracts will be second-best due to the trade-off between risk sharing and incentives. From this work, economists have followed studying the relationships between a company's shareholders (principal) and the CEO (agent) (Grossman and Hart, 1983). The classic principal-agent models establish that because the shareholders of a firm are unable to observe the effort exerted by the CEO there exists an inefficient allocation of resources. In our setup, the actions taken by the agent are also unobservable by the principal and we reach a similar situation, yet the allocation of resources changes overtime and is related strongly with the agent bargaining power.

Dynamic principal-agent models build upon this literature by presenting a way of formally studying long-term relationships between the principal and the agent, as in Spear and Srivastava (1987), and Wang (1997). This relationship is usually analyzed by means of the maximization of the shareholders' discounted expected utility subject to two constraints. The first one is the individual rationality constraint which ensures that the CEO will only accept the contract if it grants him expected discounted utility higher than or equal to his reservation utility. The second one is the incentive compatibility constraint which warrants that the CEO chooses a path of optimal effort levels that corresponds to the effort path that the shareholders want to implement. The optimal contract is achieved via two incentive devices: present and future compensation.

The trade-off between the shareholders' utility and the CEO compensation that arises from the asymmetry of information within the firm and the structure of the dynamic principal-agent model give us the opportunity to model it as a multi-objective optimization problem (Goldberg, 1989). With this approach, it is possible to model the level of bargaining power of both the agent and the principal by means of giving certain levels of priority to their corresponding utility in the objective function. Di Giannatale et al. (2021) treat the agent bargaining power as a state variable for the problem and give it an explicit dynamics to follow. We borrow from their work the same dynamics for the agent bargaining power.

Multi-objective optimization problems yield an infinite set of solutions given the lack of information about the relevance of one objective with respect to the others (Ehrgott, 2005), i.e. the real value of the bargaining power. This set of optimal solutions is called Pareto Optimal Frontier. Giving a closed-form solution to the Pareto Optimal Frontier has proven to be challenging so a good strategy to gain insight on the form of the optimal solution, and equally important the optimal set of present and future compensations, is to design and implement a computational algorithm as in Wang (1997) and Clementi et al. (2010).

Finally, this work also relates to the empirical work of identifying the source of heterogeneity in CEO compensation packages. While there exist real-world evidence of this heterogeneity, see for example Tervio (2008), Gabaix and Landier (2008), the existing theory offers mixed results when predicting CEOs compensations and efforts. For instance, Murphy and Sandino (2010) find that CEOs have acquired a higher ability to extract rent from shareholders which they argue could be a reason for this heterogeneity. However, Hölmstrom (2004) finds that there is growing pressure to diminish executives' power within firms, which makes the arguments made by Murphy and Sandino (2010) difficult to believe. Another explanation, a little more widely accepted, is that firm size is a determinant factor for the CEO compensation heterogeneity (Gabaix and Landier, 2008, Edmans and Gabaix, 2016), arguing that bigger firms have higher willingness to pay for managerial talent. Yet, Gayle and Miller (2009) find empirical evidence that links the agency costs in the U.S. with exogenous growth in firm size while not being able to link this growth with changes in managers' risk preferences.

Di Giannatale et al. (2021) find that the dynamics that the CEO's bargaining power follows could explain the difference in compensation for CEOs with even the same risk aversion parameter. This work follows the former article by combining the numerical and theoretical results of our analysis to give a better understanding of the role that within and between company factors play on determining the size and evolution of CEO compensations. Specifically, we focus on the impact of using different distributions for the output distribution, namely the Pareto Distribution which has been used to model income and wealth distributions (Zipf, 1949, Mandelbrot, 1960, Arnold, 2014) and could potentially explain salary heterogeneity at industry level.

3 The Model

In this section, we summarize the multi-objective dynamic principal-agent model proposed in Di Giannatale et al. (2021).They propose a discrete-time model in which two individuals engage in a contractual relationship from time $t = 0$ until infinity: a risk neutral principal and a risk averse agent. Both seek to maximize their discounted expected utility and have a common discount rate $\beta \in (0,1)$.

The agent utility function, $v(w_t, a_t)$, is assumed to be continuous, bounded, strictly increasing with respect to w_t and a_t. Furthermore, it is strictly concave with respect to w_t, and strictly decreasing and convex with respect to a_t. In this function, $w_t \geq 0$ represents the agent's salary which is paid at the end of every period, while a_t represents the agent's effort choice made at the beginning of every period. We assume a_t is drawn from a compact set $A = [\underline{a}, \overline{a}]$, and it is unobservable to the principal.

Every period $t \geq 1$ certain output y_t, obtained from the compact set Y, is produced and observed by the principal and the agent. There is assumed to be an stochastic relationship between the output realization and the agent's effort choice which is given by the time-invariant distribution $F(y_t \mid a_t) > 0$ for all $y_t \in Y$ and for all $a_t \in A$. Di Giannatale et al. (2021) also assume that this distribution has a density f and that the distribution of outputs is *i.i.d.* from period to period, for a given action.

As in Di Giannatale et al. (2021), we assume that at the beginning of the principal-agent relationship, at $t = 0$, the agent has a bargaining power given by δ_0, where $\delta_0 \in [0,1]$; while the principal has a bargaining power of $(1 - \delta_0)$. It is also assumed that the principal and the agent had reached an agreement, at $t = 0$, about the law that governs the movement of the agent's bargaining power period by period, $\delta_t = z(\delta_{t-1}, y_t)$. So, the contract that defines the infinite relationship of the principal and the agent follows this timeline: At $t = 1$, given a value of $\delta_0 \in [0,1]$, the agent decides $a_1(\delta_0) \in A$, at the end of the period output $y_1(\delta_0) = y_1(a_1(\delta_0))$ is drawn from the distribution $F(y \mid a_1(\delta_0))$, and the agent receives a compensation $w_1(y_1(\delta_0))$. The principal receives $y_1(\delta_0) - w_1(y_1(\delta_0))$, and the agent's bargaining power for $t = 2$, $\delta_1 = z(\delta_0, y_1) \in [0,1]$ is defined by the agreed-upon law of motion. Now assume that the principal and the agent employ history-dependent pure strategies and that the game is repeated from $t = 2, 3, ...$

The former mentioned relationship ensures that at any time t there is a history of output realizations $h^t = \{(\delta_s, y_{s+1}(\delta_s))\}_{s=0}^{t-1}$; with $h^0 = \delta_0$, such that $y_{s+1}(\delta_s) \in Y$ and $\delta_s \in [0,1]$ for all $s = 1, 2, \ldots, t-1$. The principal's decision is $w_t(h^t)$, and the agent's decision is $a_t(h^{t-1})$, because the effort decision has to be made before y_t has been realized and given the value of $\delta_t(h^{t-1})$ at the beginning of the period. Let $\pi(h^{t+\tau}; h^t, a_t)$ be the probability distribution of $h^{t+\tau}$ conditional on h^t and a_t. This distribution is recursively expressed in the following way:

$$d\pi(h^{t+\tau} \mid h^t, a_t) = f(y_{t+\tau} \mid a(h^{t+\tau-1}))d\pi(h^{t+\tau-1} \mid h^t, a_t),$$

with

$$d\pi(h^{t+1} \mid h^t, a_t) = f(y_{t+1} \mid a(h^t)).$$

The value functions that the principal and the agent, respectively, derive from the sub-game starting from h^t are given by:

$$\begin{aligned} U(h^t, w, a) &= \sum_{\tau=0}^{\infty} \beta^\tau \int_Y \left[y_{t+\tau} - w(h^{t+\tau}) \right] d\pi(h^{t+\tau} \mid h^t, a); \\ V(h^t, w, a) &= \sum_{\tau=0}^{\infty} \beta^\tau \int_Y v(w(h^{t+\tau}), a(h^{t+\tau-1})) d\pi(h^{t+\tau} \mid h^t, a). \end{aligned}$$

Given sequences $\delta_t = \{\delta_t(h^{t-1})\}$ and $w_t = \{w_t(h^t)\}$, the sequence $a_t = \{a_t(h^{t-1})\}$ is incentive compatible at h^t if:

$$V(h^t, w, a) \geq V(h^t, w, \overline{a}) = \sum_{\tau=0}^{\infty} \beta^\tau \int_Y v(w_t(h^{t+\tau}), \overline{a}(h^{t+\tau-1})) d\overline{\pi}(h^{t+\tau}; h^t, \overline{a}),$$

for any other sequence $\overline{a}_t = \{\overline{a}_t(h^{t-1})\}$, and $\overline{\pi}$ is the distribution in the future histories induced by δ_t, y_t, w_t and $\overline{a}_t$.

A contract $\phi_t^{\delta_0}$ is defined by a history-dependent agent's effort recommendation $a_t(h^{t-1})$, and a history-dependent agent's compensation plan $w_t(h^t)$. The agent's history-dependent bargaining power values $\delta_{t+1}(h^t)$ are determined by the agreed-upon law of movement. That is, a contract is given by:

$$\phi_t^{\delta_0} = \left\{a_t(h^{t-1}), w_t(h^t)\right\}.$$

We say that a contract $\phi_t^{\delta_0}$ is feasible if:

$$a_t(h^{t-1}) \in A \qquad \forall h^{t-1} \in ([0,1] \times Y)^{t-1}, \quad \forall t \geq 1; \tag{1}$$

$$0 \leq w_t(h^t) \leq y_t \qquad \forall h^t \in ([0,1] \times Y)^t, \quad \forall t \geq 1; \tag{2}$$

and also the agreed-upon law of motion of the agent's bargaining power must hold:

$$\delta_{t+1}(h^t) = z(h^t) \in [0,1] \qquad \forall h^t \in ([0,1] \times Y)^t, \quad \forall t \geq 0. \tag{3}$$

Condition (1) ensures that the agent's efforts belong to the set of admissible effort values. Condition (2) is the capacity restriction and requires the agent's salary to be not greater than the current output. Condition (3) requires that any value of the agent's bargaining power is drawn from the law of motion in the interval $[0,1]$.

For any given δ_t, two conflicting objective functions are optimized: the *ex-ante* agent's discounted expected utility, and the *ex-ante* principal's discounted expected utility, subject to incentive compatibility and feasibility. The solution to this maximization problem is not unique, but a series of contracts that satisfy Pareto optimality. A contract $\phi_t^{\delta_0}$ is Pareto optimal if there is no other feasible and incentive compatible contract $\varphi_t^{\delta_0}$ such that $(U(h^t, \varphi_t^{\delta_0}), V(h^t, \varphi_t^{\delta_0})) \succeq (U(h^t, \phi_t^{\delta_0}), V(h^t, \phi_t^{\delta_0}))$, for all h^t (Di Giannatale et al., 2021).

The dynamic problem is now transformed into a static variational one as in Spear and Srivastava (1987). The continuation profile from time $t+1$ onwards for contract ϕ_t^{δ} at any t, where

δ is the initial bargaining power of the agent, given h^t, is determined by $\phi_t^\delta \mid h^t$. This implies a continuation value from time $t+1$ onwards of $U(\phi_t^\delta \mid h^t)$ for the principal, and of $V(\phi_t^\delta \mid h^t)$ for the agent.

A contract ϕ_t^δ is temporary incentive compatible if, for all t and for all h^t:

$$a_t(h^{t-1}) \in \arg\max_{a\in A} \int_Y [v(w_t(h^t), a) + \beta V(\phi_t^\delta \mid h^t)] f(y_t; a) dy_t. \tag{4}$$

This constraint ensures that there will be no deviations in the optimal path of the agent's effort decisions, for any δ_t.

For every $\delta \in [0,1]$, define $\mathscr{W}(\delta)$ as the set of the principal's and the agent's discounted expected utility values that are generated by contracts that are feasible, incentive compatible, and characterized by the agent's initial bargaining power given by δ and the agreed-upon law of motion z, as follows:

$$\mathscr{W}(\delta) = \{(U(\phi^\delta \mid h^0), V(\phi^\delta \mid h^0)) \mid \exists\ \phi^\delta\ \ s.t.\ \ (1),(2),(3),\ \ and\ \ (4)\}.$$

Di Giannatale et al. (2021) prove that $\mathscr{W}(\delta)$ is compact for all δ and characterize the unique series of Pareto optimal contracts in a Bellman equation. For all $(U(\delta), V(\delta)) \in \mathscr{W}(\delta)$, the Bellman Equation, yields:

$$\Gamma(U,V)(\delta) = \max_{w(\delta,y),\overline{V}(\delta,y),\overline{U}(\delta,y)} \{U(\delta), V(\delta)\}$$

where:

$$U(\delta) = \int_Y [y - w(\delta,y) + \beta \overline{U}(\delta,y)] f(y; a^*(\delta)) dy,$$

$$V(\delta) = \int_Y [v(w(\delta,y), a^*(\delta)) + \beta \overline{V}(\delta,y)] f(y; a^*(\delta)) dy.$$

Subject to

$$a^*(\delta) \in \arg\max_{a(\delta)\in A} \int_Y [v(w(\delta,y),a(\delta)) + \beta \overline{V}(\delta,y)] f(y;a(\delta))dy; \tag{5}$$

$$0 \leq w(\delta,y) \leq y \quad \forall \; y \in Y; \tag{6}$$

$$\delta' = z(\delta,y) \in [0,1]; \tag{7}$$

$$(\overline{U}(\delta,y), \overline{V}(\delta,y)) \in \mathscr{W}(\delta) \quad \forall \; y \in Y. \tag{8}$$

where (5) is the incentive compatibility constraint; (6) indicates the agent's temporary inability to borrow; (7) guarantees that the future vale of the agent's bargaining power belongs to the interval $[0,1]$, and (8) ensures that the principal's and the agent's future utility plans are feasible. Di Giannatale et al. (2021) prove that $(U^*(\delta),V^*(\delta))$ is a fixed point of Γ, i.e. $(U^*(\delta),V^*(\delta)) = \Gamma(U^*,V^*)(\delta), \quad \forall \delta \in [0,1]$.

The operator Γ satisfies Blackwell's sufficient conditions for a contraction, and the contraction mapping theorem ensures that the fixed point $(U^*(\delta),V^*(\delta))$ is unique for all $(U,V) \in \mathscr{W}(\delta)$. This means that along the resulting Pareto Frontier, PF^*, there exists only one pair if maximal values of the principal's and the agent's discounted expected utilities, given a value of δ, for all $(U,V) \in \mathscr{W}(\delta)$; and vice-versa. Now, PF^* must be non-increasing because otherwise either the principal or the agent can achieve a higher level of discounted expected utility and the other individual would be better off (Spear and Srivastava, 1987).

According to Hernández-Lerma and Romera (2004), this multi-objective dynamic optimization problem admits the following Pareto Weights representation (PWR) with δ as the state variable:

$$\max_{w(\delta,y)} [\delta V(\delta) + (1-\delta)U(\delta)] \tag{9}$$

subject to constraints (5), (6), (7), and (8). Notice that each objective function has a level of priority associated; that is, δ is the priority assigned to the *ex-ante* agent's discounted expected utility, and $1-\delta$ is the priority assigned to the *ex-ante* principal's discounted expected utility.

Di Giannatale et al. (2021) define an algorithm to numerically find the stationary point in the Bellman equation in the case where the agent effort and the firm output are drawn from discrete sets. The numerical characterization of the problem and the results drawn from the algorithm allow a first approach to an empirical identification of the agent bargaining power dynamics along the contractual periods. In the next sections, we expand on this idea by numerically solving different parameterizations of the problem. Then, we use these results to propose a more robust empirical model for the identification of the bargaining power dynamics.

4 Numerical Approach

In this section, we present the computational strategy, which is similar to that in Di Giannatale et al. (2021), to numerically approximate the optimal solutions of a parameterized version of the multi-objective dynamic principal-agent model proposed in the previous section. First, we will revise the Pareto Family of Probability Distributions. With this in mind, we specify the functional forms and parameter values we use in our computational program, then we will present the algorithms that outline our computational strategy, and finally we present the results for the numerical exercises.

4.1 Generalized Pareto Family of Distributions

This subsection describes the Bounded Generalized Pareto Family of probability distributions and the properties it has that make it suitable to our model. The Pareto distribution was first introduced by Pareto (1964) as a way to model income and wealth distributions in European countries. Pareto observed that in many populations the number of individuals in the population whose income exceeded a given level x was well approximated by $Cx^{-\alpha}$, for some real numbers C and $\alpha > 0$. This approximation only worked for large values of x.

Despite criticisms, it became generally accepted that most income distributions did indeed exhibit this type of tail behaviour, later called *Paretian tail behavior* (Arnold, 2014). Paretian tail behavior ensures that the tail values of the distribution have a relatively large probability to be observed, contrasting with classical distributions as the standard normal distribution. This behaviour makes the Pareto Distribution suitable to model output distributions across firms.

Formally, we characterize a random variable X with a Pareto Distribution by its survival function as follows:

Definition 1 (Survival function) If X is a random variable with a Pareto Distribution then the probability that X is greater than some real number x, *i.e.* the survival function, is given by:

$$\overline{F}(x) = Pr(X > x) = \begin{cases} \left(\frac{\sigma}{x}\right)^{\alpha} & x \geq \sigma \\ 1 & x < \sigma \end{cases}$$

where the *scale parameter* σ and the *shape parameter* α are strictly positive.

Definition 2 (Cumulative distribution function) The cumulative distribution function (CDF) of a Pareto random variable with parameters σ and α is

$$F_X(x) = \begin{cases} 1 - \left(\frac{\sigma}{x}\right)^{\alpha} & x \geq \sigma \\ 0 & x < \sigma \end{cases}$$

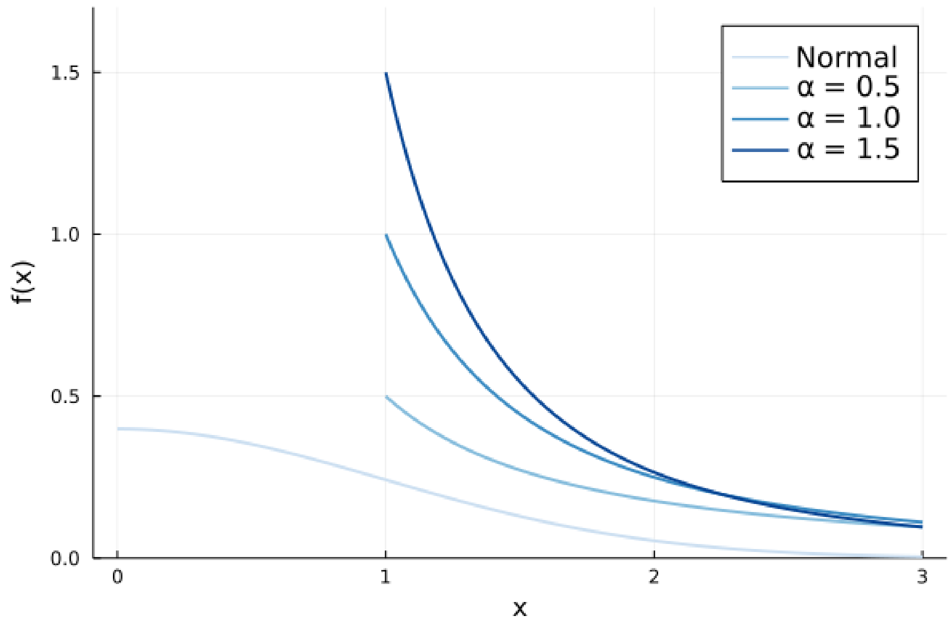
Definition 3 (Probability density function) It follows by definition 2 that the probability density function (PDF) is

$$f_X(x) = \begin{cases} \frac{\alpha\sigma^{\alpha}}{x^{\alpha+1}} & x \geq \sigma \\ 0 & x < \sigma \end{cases}$$

Figure 2 shows the the Pareto probability density function for parameter values $\sigma = 1$ and $\alpha \in \{0.5, 1.0, 1.5\}$, and the probability density function of a standard normal random variable. We observe that for lower values of α the tail of the distribution becomes *fatter* in the sense that higher values of X are more likely to be observed. This also illustrates the fact that the tail values, *i.e.* the higher values, of the random variable have a significantly positive probability to occur which is relatively higher to that in the standard normal pdf.

In order to use the Pareto distribution to model the firm output, y, we need to use a generalized bounded version of this distribution first introduced in Pickands III (1975), which allows for the random variable X to be drawn from a compact set. The probability density function of this generalized bounded Pareto distribution is given by:

Figure 2: Pareto PDF comparison, $\sigma = 1$.



Source: Own Elaboration.

$$f_X(x;\alpha) = \frac{\alpha L^{\alpha} x^{-\alpha-1}}{1-\left(\frac{L}{H}\right)^{\alpha}}; \tag{10}$$

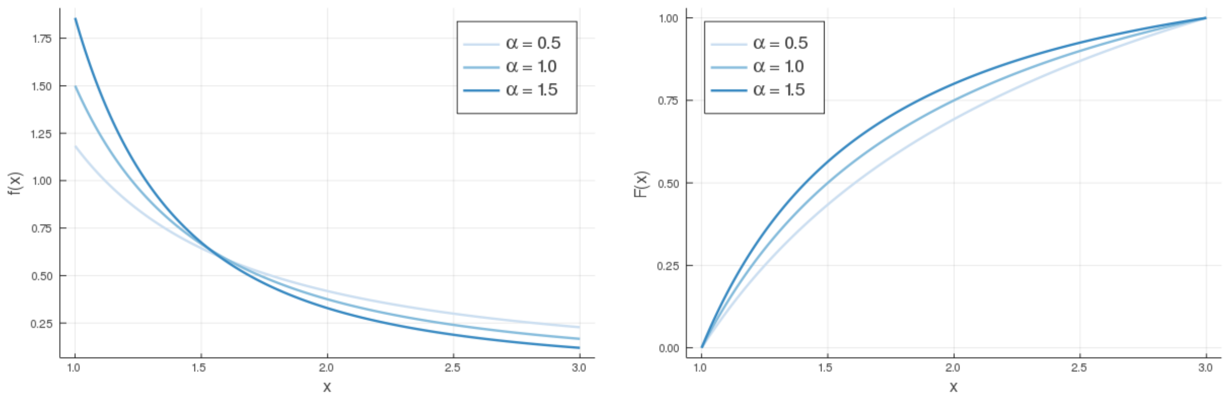
where $L(H)$ denotes the minimal (maximal) value for X, and $\alpha > 0$ is the shape parameter. From equation (10) we get that the cumulative distribution function of the bounded Pareto distribution is given by:

$$F_X(x;\alpha) = \frac{1-L^{\alpha} x^{-\alpha}}{1-\left(\frac{L}{H}\right)^{\alpha}}. \tag{11}$$

Figure 3 shows the probability and cumulative density functions of a bounded Pareto distribution with parameters $L = 1, H = 3$, and $\alpha \in \{0.5, 1.0, 1.5\}$. The panel (a) of Figure 3 illustrates that, similarly to the standard Pareto distribution, for lower values of α the probability of observing higher values of X increases. We use this fact to use α as the agent effort to solve the model numerically. We do this by letting $Y = \{y_L, y_H\}$ and $P_H(a) = P(y = y_H|a) = F(x;a), P_L(a) = P(y = y_L|a) = 1 - P_H(a)$ where x is the midpoint between L and H.

Notice that with this characterization the expected utility values for both the agent and the principal become weighted sums instead of integrals in our maximization problem. We follow

Figure 3: Bounded Pareto Distribution.



(a) Bounded PDF. **(b)** Bounded CDF.

Source: Own Elaboration.

Rogerson (1985) and use the first-order approach to solve this discrete version of the problem subject to condition (5), which requires the agent's effort to be incentive compatible. In order to use this approach, the bounded Pareto distribution should fulfill two conditions. The first is to satisfy the monotone likelihood-ratio condition (MLRC).

Definition 4 The functions $\{P_L(a), P_H(a)\}$ are said to satisfy the MLRC if $a \leq a'$ implies that $P_L(a)/P_L(a') \geq P_H(a)/P_H(a')$.

The MLRC implies that the observation of a higher level of output allows the statistical inference that the agent worked harder in the sense of stochastic dominance, this also implies that increases in effort cause output to increase in the same sense.

Proposition 1. $\{P_L(a), P_H(a)\}$ *as defined above satisfy the MLRC.*[2]

The second condition that the probability functions need to satisfy is the convexity of the distribution function condition (CDFC).

Definition 5 The functions $\{P_L(a), P_H(a)\}$ satisfy the CDFC if $P_L''(a)$ and $P_L''(a) + P_H''(a)$ are non-negative.

[2] The proof is in Appendix A.1.

By MLRC, the probability of a low outcome decreases as the agent works harder. The CDFC requires that the function decreases at a decreasing rate.

Proposition 2. *The functions* $\{P_L(a), P_H(a)\}$ *satisfy the CDFC.*[3]

Propositions 1 and 2 allow us to relax constraint 5 for the first order condition of the maximization program. With our current characterization of the problem, and assuming $v(w(\delta,y),a(\delta))$ is additively separable in the sense that $v(w(\delta,y),a(\delta)) = \nu(w(\delta,y)) - a(\delta)$, this condition is given by

$$\frac{\partial F(x;a(\delta))}{\partial a}\left[\left(\nu(w(\delta,y_H)) + \beta\overline{V}(\delta,y_H)\right) - \left(\nu(w(\delta,y_L)) + \beta\overline{V}(\delta,y_L)\right)\right] - 1 = 0. \qquad (12)$$

4.2 Functional Forms and Parameter Values

In this sub-section, we enlist a series of assumptions regarding functional forms and parameter values we made with the purpose of implementing the computational strategy.

The principal's utility function for one period consumption is given by: $u(y,w(y,\delta)) = y - w(y,\delta)$. That of the agent is given by: $v(w(y,\delta),a(\delta)) = \frac{w^{1-h}(y,\delta)}{1-h} - a(\delta)$, where $0 < h < 1$. Notice that it is of the CRRA type with respect to current salary w, and that the coefficient of relative risk aversion is given by h, where higher values of h imply higher degrees of relative risk aversion. For the benchmark result, the value of $h = 0.5$. The agent and the principal discount the future at a rate of $\beta = 0.96$.

The agent's feasible effort choices are continuous and belong to the set $A = [a_L, a_H]$, where $a_L = 0.1$ is the lowest effort choice and $a_H = 1$ is an upper bound for the agent effort. Also, there are two levels of output: low (L) or high (H), described by the set $Y = \{y_L, y_H\}$. We take $y_L = 0.4$ and $y_H = 0.8$ as the benchmark parameters. The probability functions that formalize the stochastic relationship between effort and output are $P_H(a)$ and $P_L(a)$ as described in the previous sub-section.

[3] The proof is in Appendix A.2.

We take the law of motion for the agent's bargaining power from Di Giannatale et al. (2021) as follows:

$$\delta' = z(\delta, y) = \begin{cases} \min\{1, \delta + \varepsilon \cdot \frac{y}{y_H}\} & if \quad y = y_H, \\ \max\{0, \delta - \varepsilon \cdot \frac{y}{y_H}\} & if \quad y = y_L. \end{cases}$$

where ε is an arbitrarily small and positive number. This law of motion provides incentives in the form of a greater next-period bargaining power if output y_H is observed, and a punishment in the opposite direction if output y_L occurs at the current period. According to the authors, the parameter ε can be seen as a measure of how closely future values of the CEO's bargaining power showcase rewards or punishments for good versus bad performances of the firm. We use $\varepsilon = 0.001$ as a reference in our analysis.

4.3 Computational Algorithm

We now present the algorithm that outlines the computational program we designed to obtain a numerical solution of (9).[4] The process of finding this solution requires two steps. The first step is to find the set of admissible values for our state variable i,e., finding the feasible agent bargaining powers. We find the minimal admissible values of the state variable, δ_{min} , and the maximal admissible value of this variable, δ_{max}, via the algorithm in table 2. Once the algorithm finds these values, we discretize the set of admissible bargaining powers and start the second step of the process. The algorithm in table 3 presents the recursive process of finding the stationary solution of the Bellman equation for the Pareto Weights representation of our multi-objective dynamic principal-agent model.

4.4 Numerical Results

In this subsection we present the main results of our numerical implementation. In order to highlight the impact of taking into account both the bargaining power dynamics and the Pareto distribution in our analysis, we compare the results with those of some benchmark models. The first benchmark model is the standard dynamic principal-agent model (SDPA), the second refer-

[4] For more details on the algorithm see the electronic appendix B.

Table 2: Step 1 of the Algorithm to find the stationary solution.

Objective: Find the set $[\delta_{min}, \delta_{max}]$ of feasible bargaining powers.
Initialization: Set $\delta = 0$, $V_t(\delta) = 0_N$, $U_t(\delta) = 0_N$, $\Delta = 0.005$ and $K = 1$. *Step 1:* Solve for (w_h, w_l) in (9), given $\delta_l = z(\delta, y_l)$ and $\delta_h = z(\delta, y_h)$ subject to incentive constraint, save optimal utility values V^* and U^*. *Step 2:* Set $\delta' = \delta + \Delta$. *Step 3:* Solve for (w_h, w_l) in (9) given δ', save the corresponding optimal utility values as V^{**} and U^{**}. *Step 4:* If $V^* \neq V^{**}$ or $U^* \neq U^{**}$, set $\delta_m in = \delta$ and go to *step 5*. Otherwise, set $\delta = \delta'$ and go to *step 2*. *Step 5* Set $\delta = 1$ *Step 6:* Solve for (w_h, w_l) in (9) given $\delta_l = z(\delta, y_l)$ and $\delta_h = z(\delta, y_h)$ subject to (12), save the corresponding optimal utility values as V^* and U^*. *Step 7:* Set $\delta' = \delta - \Delta$. *Step 8:* Solve for (w_h, w_l) in (9) given δ', save the corresponding optimal utility values as V^{**} and U^{**}. *Step 9:* If $V^* \neq V^{**}$ or $U^* \neq U^{**}$, set $\delta_m ax = \delta$ and stop. Otherwise, set $\delta = \delta'$ and go to *step 7*.

Source: Own Elaboration.

ence is the multi-objective static principal-agent model (MOSPA), finally the third reference is the multi-objective dynamic principal-agent model with no dynamics for the agent's bargaining power (MODPA1).[5] We refer to the model described in section 3 as (MODPA2).

4.5 Stationary Results

First, in Figure 4 we show a comparison between the Pareto frontiers of the benchmark models and the model described in section 3.[6] Panel (a) of Figure 4 shows that the Pareto frontier from the SPDA, MOSPA, and MODPA1 models coincide. This depicts that first, in this case there is no loss of efficiency when going from a static to a dynamic setting. And second, that solving the multi-objective optimization problem (MODPA1) is in fact equivalent to solving the SSPA model. In Panel (b) of Figure 4, we show that for a small value of ε, the Pareto frontier of our

[5] See Appendix A.3 for more details about these models.
[6] Reefer to Appendix B for the solution data.

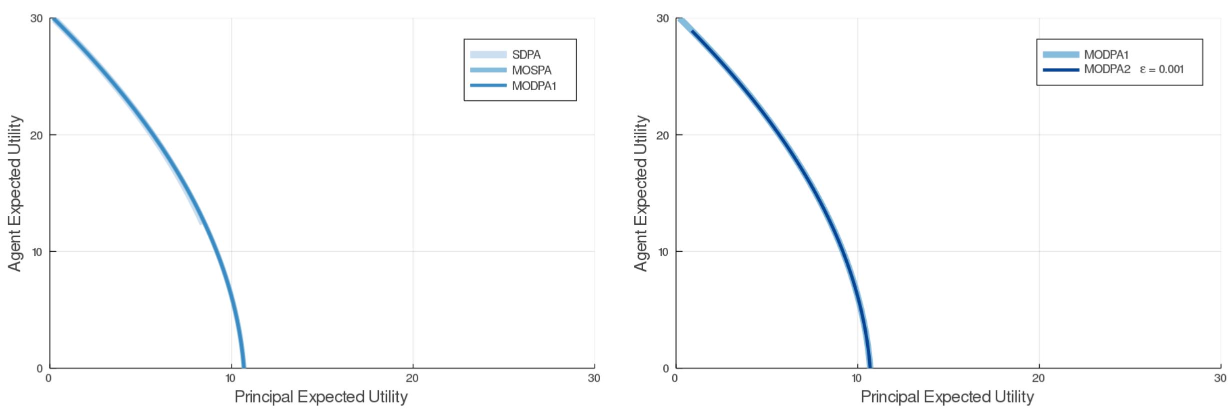
Table 3: Step 2 of the Algorithm to find the stationary solution.

Objective: Solve program (9) subject to constrains (6), (7), (8) and (12). Given a set $[\delta_{min}, \delta_{max}]$ of feasible bargaining powers.
Initialization: Set $N = ((\delta_{max} - \delta_{min}) \times 2)/\varepsilon + 1$, $\delta = \delta_{min}$, $V_t(\delta) = 0_N$, $U_t(\delta) = 0_N$ and $K = 1$. *Step 1:* Solve for (w_h, w_l) in (9) given $\delta_l = z(\delta, y_l)$ and $\delta_h = z(\delta, y_h)$ subject to (12), save the corresponding optimal value as S^*. Set $V_{t+1}[K], U_{t+1}[K]$ as the corresponding agent and principal expected utilities in S^*. *Step 2:* If $\delta < \delta_{max}$ set $\delta' = z(\delta, y_h)$, $K' = K + 1$ and go to *step 1* using δ', K'. If $\delta = \delta_{max}$ go to *step 3*. *Step 3:* If $V_t(\delta) = V_{t+1}(\delta)$ and $U_t(\delta) = U_{t+1}(\delta)$, stop. Otherwise, set $\delta' = \delta_{min}$, $K = 1$ and go to *step 1* using δ', K'.

Source: Own Elaboration.

model actually converges to that of the MODPA1. This is expected given that when $\varepsilon \to 0$ in the proposed law of motion, $\delta' \to \delta$, *i.e.*, the bargaining power will remain constant through the contractual relationship as in MODPA1. All in all, Figure 4 serves as a way to show consistency of the numerical solution of MODPA2.

Figure 4: Pareto Frontier: Model Comparison.



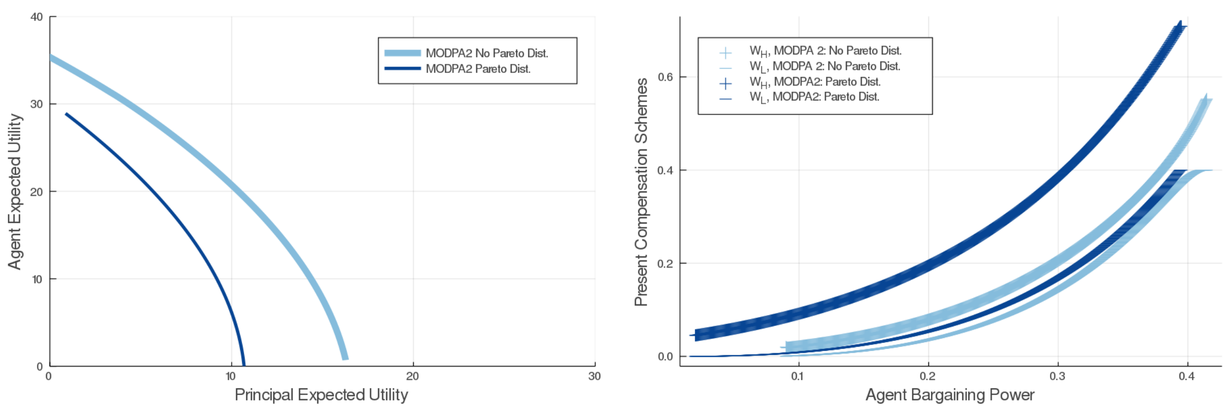
(a) Single objective and multi-objective approach. **(b)** Bargaining dynamics and no dynamics.

Source: Own Elaboration with data obtained via numeric algorithm.

Figure 5 shows the results from solving the MODPA2 program with an output distribution drawn from a Pareto PDF and a more conventional output distribution: we draw from Wang (1997) the output distribution on a discrete action choice model given by $f(y_H|a_L) = f(y_L|a_H) =$

1/3, $f(y_H|a_H) = f(y_L|a_L) = 2/3$. The Pareto frontier shown in panel (a) of Figure 5 depicts a loss of welfare when considering the Pareto distribution of outputs. Panel (b) also suggest a lesser degree of insurance for the agent when considering the Pareto distribution of output. It is also worth noting that considering a fat tail distribution implies higher probability of higher outputs even when the agent exerts low effort. So, in our exercise when considering Wang's distribution of output, the agent exerts high effort ($a = 0.2$) optimally. But when considering the Pareto distribution, the agent exerts low effort ($a = 0.1$). This implies a higher effort productivity when considering the Pareto distribution. All in all, results from Figure 5 can be interpreted as follows: when facing higher probabilities of high output, the agent will be less inclined to exert effort which in turns leads to a lesser degree of insurance and a loss of overall welfare.

Figure 5: Pareto PDF Effect on Results.



(a) MODPA2: Pareto frontier.

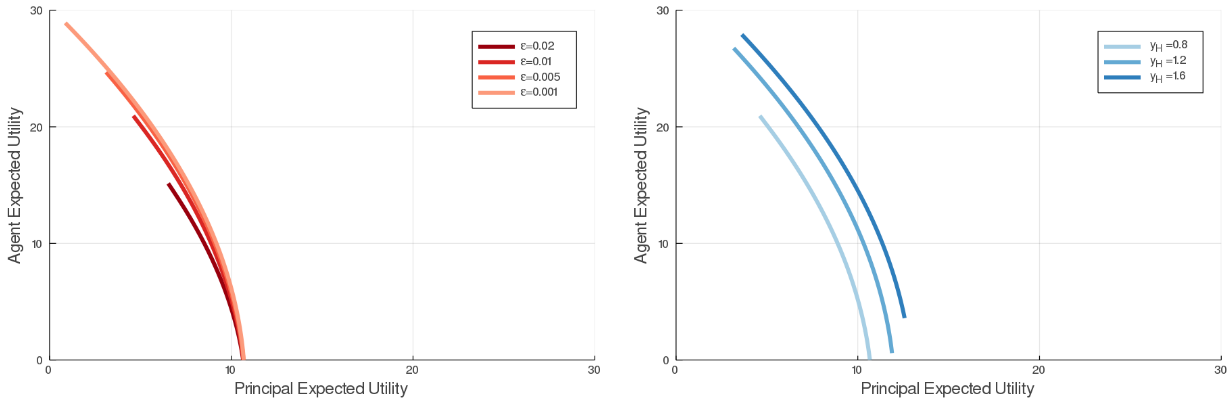
(b) MODPA2: present compensations.

Source: Own Elaboration with data obtained via numeric algorithm.

Figures 6, 7, and 8 show results from comparative statics exercises with different values for the parameters of the law of motion for the agent bargaining power and the probability density function for the company output. The parameter ε takes values from 0.001 to 0.02 in our comparison, giving different degrees of movement to the agent bargaining power. We examine the role of the Pareto distribution in our results by means of changing its shape, *i.e*, we expand the upper bound, y_H, of the distribution from 0.8 to 1.2 and to 1.6.

The results from panel (a) of Figure 6 show that there is a slight lost of welfare in the form

Figure 6: Pareto Frontier: Parameter Comparison.

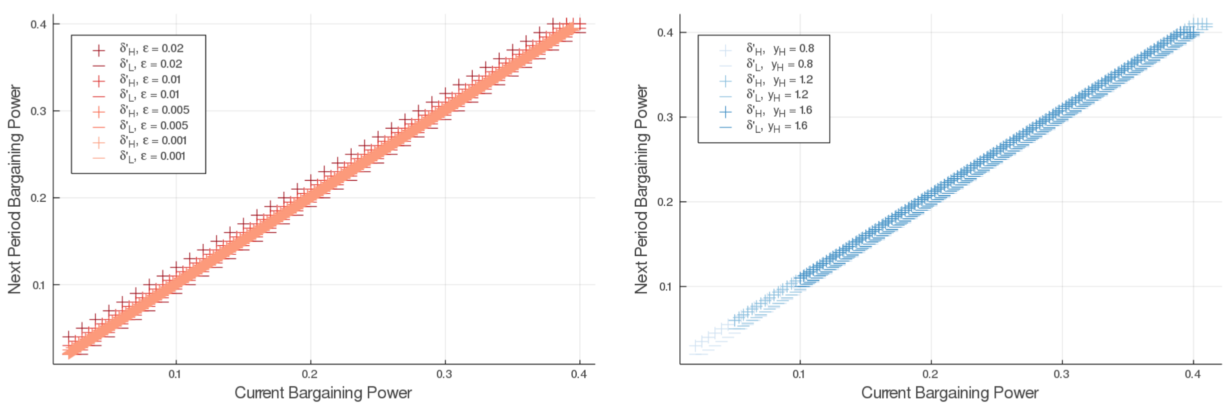


(a) Different Values for ε. **(b)** Different Pareto Shape Parameters.

Source: Own Elaboration with data obtained via numeric algorithm.

of Pareto-dominated solutions in the Pareto Frontier as ε increases. Notice how for the risk adverse agent this loss of welfare is accompanied with lower levels of maximum expected utility as greater ε values bring higher uncertainty about next period bargaining power, effectively diminishing the expected utility for the agent. As the principal is risk neutral, his maximum expected utility stays relatively the same. In panel (b) of this Figure 6, we show that the possibility of greater expected outputs as the Pareto PDF gets a fatter tail brings greater welfare for both parties of the firm.

Figure 7: Agent's Promised Bargaining Power: Parameter Comparison.



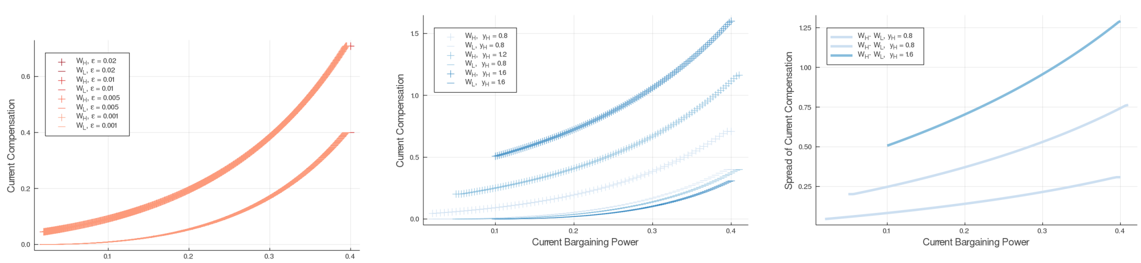
(a) Different Values for ε. **(b)** Different Pareto Shape Parameters.

Source: Own Elaboration with data obtained via numeric algorithm.

Figure 7 shows the changes in promised bargaining power for high and low output realizations when the parameters of the model change. Notice that the relationship between current and next period bargaining power is linear for both high and low output realization. The variations in ε depicted in panel (a) have an impact in the distance between the agent's promised bargaining power results for high and low output realizations. As ε increases, the level of insurance that the agent receives decreases, which is in line with the results shown in Figure 6. In panel (b) we show that the variation in the shape of the output distribution does not affect next period bargaining power results for high and low output realizations.

In Figure 8 we present the changes in agent compensation for both low and high output realizations for different parameters. In panel (a), as usual, we show the variations on the parameter ε. We find that this parameter does not have an impact on present compensation, showing that the dynamics of future bargaining power do not affect current salary for the CEO. Panel (b) show the corresponding comparative statics when the output distribution varies. Notice that a fattier distribution brings a higher spread between current salary in the case of a high and a low output distribution, as shown in panel (c). It is also worth noticing that even in the case where the lower bound for the distribution is fixed, lower salaries in the case of low output come with fattier tail distributions.

Figure 8: Current Compensation: Parameter Comparison.



(a) Different Values for ε.

(b) Different Pareto Shape Parameters.

(c) Spread of Compensations in (b).

Source: Own Elaboration with data obtained via numeric algorithm.

From the results illustrated in Figures 7 and 8, we conclude that the initial value of the CEO's bargaining power and the value of ε in the law of motion determine the evolution of the optimal contractual arrangements between the CEO and the shareholders. This is consistent with the results found in Di Giannatale et al. (2021). Our results contribute with the observation that the shape of the output distribution dictates the level of present compensation for the CEO but does not dictate the evolution of the contract.

4.6 Simulation Results

In this subsection we present the results from a series of Monte Carlo simulations of contractual relationships. We aim to answer how does the shape parameter of the Pareto distribution and the dynamics of the bargaining power affect the evolution of the relationship.

Table 4: Monte Carlo Estimations.

	Conventional PDF			Pareto PDF			Different Means T-test
Variable	Max δ	Mean	St. Dev.	Max δ	Mean	St. Dev.	P-Value
(1) Periods Until Max Bargaining Power	0.42	560.08	33.46	0.395	500.92	42.45	$< 1e-99$
(2) Total Discounted Compensation	0.42	16.41	0.67	0.395	15.65	0.71	$< 1e-99$

Source: Own Elaboration with data obtained via numeric simulations.

First, in the first row of Table 4 we show Monte Carlo estimations for the mean and standard deviation of the expected number of contracting periods until the agent reaches for the first time his maximum bargaining power in the contract.[7] Notice that in columns 2 and 5, that the difference in output distribution implies changes in the maximum bargaining power for the agent. In particular, the Pareto distribution implies a lower maximum bargaining power for the agent. The third and fourth columns of table 4 show the estimations in the case of the conventional PDF, while columns 6 and 7 show the estimations for the benchmark Pareto distribution. The p-value of the T-test for difference in means of column 8, shows that the expected number of periods until the agent reaches his maximum bargaining power is significantly lower when the output is drawn from a Pareto distribution.

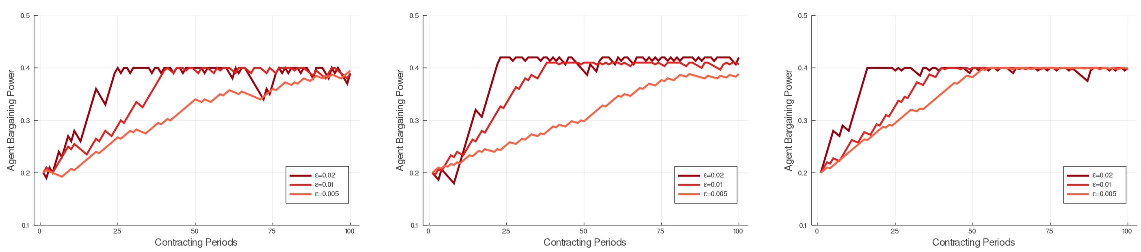
In row (2) of Table 4 we show similar estimations for the total discounted compensation received by the agent in 100 contracting periods. Columns 4 and 7 show that the standard devia-

[7] Estimations for Table 4 are based on 10,000 simulated optimal contracts. See appendix B for details.

tion of this compensation is greater when considering the Pareto distribution, while the expected compensation is lower with this distribution. This highlights the fact that using a fat tail distribution does not necessarily imply higher expected salaries, but higher heterogeneity in salaries across firms. The results from in Table 4 show a trade-off between bargaining power and present compensation for the agent when he faces a fat tail distribution of the output given his effort. He receives higher levels of bargaining power in a shorter period but he gets an overall lower total compensation due to the lack of insurance in the optimal contract when facing the Pareto distribution.

In order to further analyze the role of the bargaining dynamics when we consider the Pareto distribution, we simulate 100 contacting periods when the agent starts with a bargaining power of $\delta_0 = 0.20$; when the value of ε in the bargaining dynamics takes values from $\{0.02, 0.01, 0.005\}$, the low outcome is set to 0.4 and high outcomes possibilities are drawn from $\{0.8, 1.2, 1.6\}$.

Figure 9: Simulation: Agent's Bargaining Power.



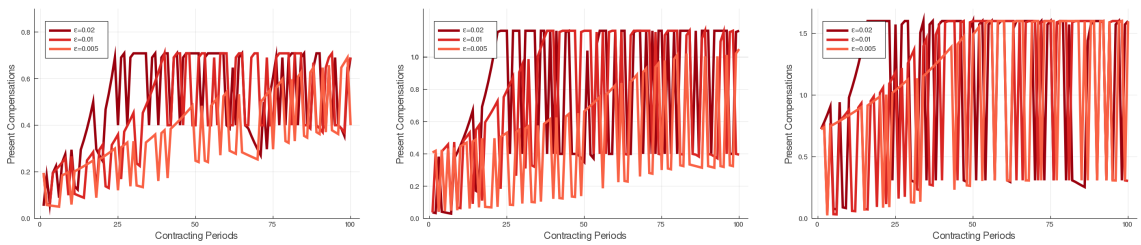
(a) $y_h = 0.8$. **(b)** $y_h = 1.2$. **(c)** $y_h = 1.6$.

Source: Own Elaboration with data obtained via numeric simulation.

In Figure 9 we show the results of the agent's bargaining power, which has a positive trend with respect to the contractual periods. This positive trend is explained by the higher probability of the higher output, given the optimal effort $(a = 0.1)$. Effort productivity, meaning higher probability of higher outcomes when exerting the same effort, also has a positive correlation with bargaining power increments. The greater the parameter ε is, the faster the agent's bargaining power reaches its maximal value, i.e. the more variability between present and next period bargaining power, the faster the CEO can reach higher levels of bargaining power if he performs properly in the company.

Figure 9 also shows that for all three sets of parameters for the high and low outputs there is not a significant difference in the maximum bargaining power reached by the agent. So, a higher probability of a higher output does not imply higher bargaining powers achievable by the CEO.

Figure 10: Simulation: Agent's Compensation.



(a) $y_h = 0.8$. **(b)** $y_h = 1.2$. **(c)** $y_h = 1.6$.

Source: Own Elaboration with data obtained via numeric simulation.

Finally, the simulation results regarding the agent's salary are presented in Figure 10. We observe that even for lower values of ε, there is a significant variability of salary with respect to the contractual periods, contrasting with the results in Di Giannatale et al. (2021), where for low values of ε there is a positive relationship between the agent's salary and the contractual periods. This is due to the lack of insurance provided to the agent in our results. We also observe that the higher values of ε and/or y_H, the faster the agent reaches higher values of bargaining power, and the more variability there is in his salary. This variability also implies something similar to Figure 9, higher (lower) salaries do not necessarily imply higher (lower) bargaining power for the agent.

The simulation results add to the fact that the bargaining dynamics is the determinant factor for the evolution of the contractual relationship, while the shape of the distribution determines the level of risk that the CEO faces in the company and the value of his salary. All in all, results from subsections 4.5 and 4.6 imply that industry-level qualities such as productivity distribution play an important role on determining and shaping the distribution of CEO present salaries. But once CEO salary distribution is determined, firm-level characteristics such as the CEO bargaining power are the only responsible for the way that this salaries evolve.

5 Empirical Evidence

In this section, we compare the trend and seasonal components of the time series of simulated salaries presented in subsection 4.6 with observed CEO salaries. We use the Capital IQ's Compensation Detail database which contains compensation data from around 700 different CEOs of companies based in the United States ranging from the years 2001 to 2018.[8] We filtered the CEOs with complete data from the years 2006 to 2018 and obtained a sub-sample of 32 CEOs with 12 salary observations measured in constant US dollars of 2005.

Figure 11 shows the decomposition in trend, stationary, and random components from 6 particular CEO salary time series.[9] The first plot in each subfigure shows the observed annual salaries for the corresponding CEO, which we have normalized in order to maintain a constant scale from 0 to 1. Three out of the six CEOs have salaries that have increased with time, namely those of the CEOs from BancFist, Arcelor Mittal and Walt Disney. The CEOs from Hornbeck Offshore Services and The AES Corporation are observed to have a decrease in salaries over the years. Lastly, the salary of the CEO from Qurate Retal has been maintained in a constant range. The growth or shrink tendency of these salaries is captured in the trend component of the time series which is calculated using a simple moving average and is shown in the second row of each subplot.

The seasonal component of the observed salaries is presented in the third row of the corresponding subfigure. This seasonal component is a measure of the risk that the CEO faces in the form of a difference between his expected salary based on the trend and his actual annual salary. Notice that all the CEOs, no matter their salaries' trend, increasing or decreasing, face risk within their company with peak values of seasonal components ranging from -0.1 to 0.1.

The data from CEO compensations shows that salary trends are heterogeneous even amongst companies with similar characteristics such as company size, age, or industry sector. This indicates that such outside characteristics are not strong determinants for CEO compensation and

[8] *This database is not available to the public.*
[9] All 32 time series decomposition available in Appendix B.

that we should instead focus on within firm characteristics, such as the CEO bargaining power, when studying salary trends. On the other hand, CEO compensation size is shown to increase with company size and productivity as predicted with the Pareto distribution in our model.

In Figure 12 we take the simulated salaries of an optimal contract based in our model with the benchmark parameters for both the Pareto and the conventional distribution of output ($\varepsilon = 0.01$, $y_L, y_H = 0.4, 0.8$, $\delta_0 = 0.20$) and perform the same decomposition as with the observed salaries. The trend component for both output distribution show that there is an expected growth in salaries over time thanks to the bargaining power dynamics in the model. The salaries obtained with the Pareto distribution of output show a greater seasonal component than the conventional distribution, adding to the result of greater risk sharing in the company when accounting for the fat tail distribution of output. In fact, the seasonal component in the salaries simulated using a Pareto distribution is of the same magnitude than those observed in the real salaries. While that of the salaries simulated with the conventional distribution is of lesser magnitude.

This results indicate that the predictions of the model using the bargaining power dynamics and the Pareto distribution of output offer an explanation for the heterogeneity among CEO present compensations and their different paths even between companies of similar characteristics such as size, age, performance or industry sector.

6 Conclusion

This thesis contributes to the study of the heterogeneity among CEO present salaries and growing salary trends by modeling firm and industry-level characteristics that have been observed to determine CEO compensation size. We borrow from a multi-objective dynamic principal-agent model with the CEO's initial bargaining power as the state variable in order to model the governance dynamics of a company. We also introduce a fat-tail distribution in order to capture firm productivity heterogeneity.

Our model manages to capture these three important aspects of CEO compensation: heterogeneity in value, trend, and level of risk in the company to a good degree. The results in our numerical implementation imply that the form of the productivity distribution determines the risk and compensation size that the CEO receives in the current period; but has no impact on the evolution of his future bargaining power, and therefore, does not determine the evolution of the contractual relation. On the other hand, the bargaining power dynamics and more specifically the level of reward and punishment imposed by the parameter ε does not play a role on establishing current compensation levels, but has all the faculty to shape the relationship going forward. The simulated contractual relationships using the optimal results in our model find that the shape of the output distribution specifies the level of risk that the CEO faces in the company. The heavier the tail of the distribution, the more variation between salaries which implies more risk for the CEO.

Finally, this thesis might be challenged by the lack of robust data for the empirical study or the parameter characterization in the numerical approach so the results should be understood as a glance to what the part the managerial relations as well as the state of the world play on determining the salary of the CEOs; and not as a tool to make any empirical inference about the salary trends for any given CEO.

A Appendix

A.1 Proof of Proposition 1

Proof. Let $a, a' \in \mathbb{R}^+$ such that $a \leq a'$. First, it is useful to show that $P_H(a) \leq P_H(a')$, i.e. $P'_H(a) \geq 0$ for all $a \in \mathbb{R}^+$.

In fact

$$\begin{aligned}
P'_H(a) &= \frac{\partial F_X(x;a)}{\partial a} \\
&= \frac{\partial}{\partial a}\left(\frac{1-L^\alpha x^{-\alpha}}{1-\left(\frac{L}{H}\right)^\alpha}\right) \\
&= \frac{-\left(\frac{L}{x}\right)^\alpha \log\left(\frac{L}{x}\right)\left(1-\left(\frac{L}{H}\right)^\alpha\right) + \left(\frac{L}{H}\right)^\alpha \log\left(\frac{L}{H}\right)\left(1-\left(\frac{L}{x}\right)^\alpha\right)}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2} \\
&= \frac{\left(\frac{L}{x}\right)^\alpha \log\left(\frac{x}{L}\right)\left(1-\left(\frac{L}{H}\right)^\alpha\right) - \left(\frac{L}{H}\right)^\alpha \log\left(\frac{H}{L}\right)\left(1-\left(\frac{L}{x}\right)^\alpha\right)}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2} \\
&\geq \frac{\left(\frac{L}{x}\right)^\alpha \log\left(\frac{H}{L}\right)\left(1-\left(\frac{L}{H}\right)^\alpha\right) - \left(\frac{L}{H}\right)^\alpha \log\left(\frac{H}{L}\right)\left(1-\left(\frac{L}{x}\right)^\alpha\right)}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2} \\
&= \frac{\log\left(\frac{H}{L}\right)\left[\left(\frac{L}{x}\right)^\alpha\left(1-\left(\frac{L}{H}\right)^\alpha\right) - \left(\frac{L}{H}\right)^\alpha\left(1-\left(\frac{L}{x}\right)^\alpha\right)\right]}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2}.
\end{aligned}$$

Notice that for $x \in [L, H]$ we have $\frac{L}{H} \leq \frac{L}{x} \leq 1$. So we can assure that

$$0 \leq \left(\frac{L}{H}\right)^\alpha \leq \left(\frac{L}{x}\right)^\alpha$$

and

$$1-\left(\frac{L}{x}\right)^\alpha \leq 1-\left(\frac{L}{H}\right)^\alpha.$$

The former inequalities imply

$$\left[\left(\frac{L}{x}\right)^{\alpha}\left(1-\left(\frac{L}{H}\right)^{\alpha}\right)-\left(\frac{L}{H}\right)^{\alpha}\left(1-\left(\frac{L}{x}\right)^{\alpha}\right)\right]\geq 0.$$

Then,

$$\frac{\log\left(\frac{H}{L}\right)\left[\left(\frac{L}{x}\right)^{\alpha}\left(1-\left(\frac{L}{H}\right)^{\alpha}\right)-\left(\frac{L}{H}\right)^{\alpha}\left(1-\left(\frac{L}{x}\right)^{\alpha}\right)\right]}{\left(1-\left(\frac{L}{H}\right)^{\alpha}\right)^{2}}\geq 0 \tag{13}$$

From relation (13), we have $P'_H(a) \geq 0$, thus $P_H(a) \leq P_H(a')$. This in turn implies two inequalities, first

$$P_H(a)/P_H(a') \leq 1 \tag{14}$$

and second

$$1-P_H(a') \leq 1-P_H(a) \tag{15}$$

Relation (15) implies $P_L(a') \leq P_L(a)$, so

$$P_L(a)/P_L(a') \geq 1 \tag{16}$$

Finally, from inequalities (14) and (16) we get

$$P_L(a)/P_L(a') \geq P_H(a)/P_H(a')$$

□

A.2 Proof of Proposition 2

Proof. First notice that $P_L(a)+P_H(a)=1$ for all $a\in\mathbb{R}^+$, so $P_L''(a)+P_H''(a)=0$. Now, $P_L(a)=1-F_x(x;a)$. So from the proof in Appendix A.1 we have

$$\begin{aligned} P_L'(a) &= -\frac{\partial F_X(x;a)}{\partial a} \\ &= -\frac{\left(\frac{L}{x}\right)^\alpha \log\left(\frac{x}{L}\right)\left(1-\left(\frac{L}{H}\right)^\alpha\right)-\left(\frac{L}{H}\right)^\alpha \log\left(\frac{H}{L}\right)\left(1-\left(\frac{L}{x}\right)^\alpha\right)}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2}. \end{aligned}$$

Then,

$$\begin{aligned} P_L''(a) = &-\frac{\left[\left(\frac{L}{x}\right)^\alpha \log\left(\frac{L}{x}\right)\log\left(\frac{x}{L}\right)-\left(\frac{L}{x}\right)^\alpha\left(\frac{L}{H}\right)^\alpha \log\left(\frac{L}{x}\frac{L}{H}\right)\log\left(\frac{x}{L}\right)\right]\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^4} \\ &-\frac{\left[\left(\frac{L}{H}\right)^\alpha \log\left(\frac{L}{H}\right)\log\left(\frac{H}{L}\right)-\left(\frac{L}{X}\right)^\alpha\left(\frac{L}{H}\right)^\alpha \log\left(\frac{L}{x}\frac{L}{H}\right)\log\left(\frac{H}{L}\right)\right]\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^4} \\ &+\frac{2\left(1-\left(\frac{L}{H}\right)^\alpha\right)\left(\frac{L}{H}\right)^\alpha \log\left(\frac{H}{L}\right)\left[\left(\frac{L}{x}\right)^\alpha \log\left(\frac{x}{L}\right)\left(1-\left(\frac{L}{H}\right)^\alpha\right)-\left(\frac{L}{H}\right)^\alpha \log\left(\frac{H}{L}\right)\left(1-\left(\frac{L}{x}\right)^\alpha\right)\right]}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^4} \\ = &\frac{\left[\left(\frac{L}{x}\right)^\alpha \log^2\left(\frac{x}{L}\right)-\left(\frac{L}{H}\right)^\alpha \log^2\left(\frac{H}{L}\right)\right]\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^4} \\ &+\frac{\log\left(\frac{x}{L}\frac{H}{L}\right)\left(\frac{L}{X}\right)^\alpha\left(\frac{L}{H}\right)^\alpha\left(\log\left(\frac{H}{L}\right)-\log\left(\frac{x}{L}\right)\right)\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^4} \\ &+\frac{2\left(1-\left(\frac{L}{H}\right)^\alpha\right)\left(\frac{L}{H}\right)^\alpha \log\left(\frac{H}{L}\right)}{\left(1-\left(\frac{L}{H}\right)^\alpha\right)^2}\left(\frac{\partial F_X(x;a)}{\partial a}\right). \end{aligned}$$

The first term on the former equation is non-negative because $\left(\frac{L}{x}\right)^\alpha \log^2\left(\frac{x}{L}\right)\geq\left(\frac{L}{H}\right)^\alpha \log^2\left(\frac{H}{L}\right)$. The second term of the equation is also non-negative because $\log\left(\frac{H}{L}\right)-\log\left(\frac{x}{L}\right)\geq 0$. Similarly, the third term is positive given that $\frac{\partial F_X(x;a)}{\partial a}\geq 0$. Thus, $P_L''(a)\geq 0$. Therefore, $\{P_L(a),P_H(a)\}$ satisfy the CDFC. □

A.3 Benchmark Models

We use three benchmark or reference models in our numerical implementation. Here we present them:

A.3.1 The Standard Dynamic Principal-Agent Model (SDPA)

A first reference model is the standard dynamic principal-agent model. In particular, we adapt the model of Wang (1997), which is a model formulated as the maximization of the expected discounted utility of the principal subject to the participation constraint, the incentive compatibility constraint, and feasibility constraints. The value function is the principal's expected discounted utility is given by:

$$U(\hat{V}) = \mathbb{E}[y - w(y,\hat{V}) + \beta \overline{U}(y,\hat{V})],$$

where y is the observed output, $w(y,\hat{V})$ is the compensation given the observed output y and β is the discount factor of the principal and the agent. The agent´s lifetime discounted expected utility is given by :

$$V(\hat{V}) = \mathbb{E}[v(w(y,\hat{V}), a(\hat{V})) + \beta \overline{V}(y,\hat{V})],$$

where $v(w(y,\hat{V}))$ is the temporary utility function of the agent, and $a \in A$ is the effort exerted by the agent. In addition, $\overline{V}(y,\hat{V})$ is the agent future discounted utility, which is the promised expected utility from tomorrow on. the model's state variable, $\hat{V} \in \mathbf{V}$, represent the agent's reservation utility. This is an important difference with respect to our multi-objective dynamic models given that our models' state variable is the agent's initial bargaining power.

The dynamic maximization program is:

$$\max_{w(y,\hat{V}),\overline{V}(y,\hat{V})} \mathbb{E}[y - w(y,\hat{V}) + \beta \overline{U}(y,\hat{V})]$$

subject to

$$a(\hat{V}) \in argmax_{a'} \quad \hat{V}(a', \hat{V}),$$

$$V(\hat{V}, a(\hat{V})) = \hat{V},$$

$$a(\hat{V}) \in A,$$

$$0 \leq w(y, \hat{V}) \leq y \qquad \text{for all } y,$$

$$\overline{V}(y, \hat{V}) \in \mathscr{V} \qquad \text{for all } y.$$

Let $\mathscr{V}(\hat{V})$ and $\mathscr{U}(\hat{V})$ be the set of feasible and incentive compatible expected discounted utilities of the agent and principal, respectively. Wang (1997) demonstrates that $\mathscr{U}(\hat{V})$ is compact. Therefore, by virtue of the Bellman equation, there exists a principal's maximal expected discounted utility that is feasible and incentive compatible.

A.3.2 The Multi-Objective Static Model (MOSPA)

A second reference model is the multi-objective static principal-agent model. In this setting, the static contracting problem is to choose an action $a \in A$ and a compensation scheme $w(y, \delta) \in [0, y]$, $\forall y \in Y$, to maximize the Pareto Weights function of the expected utility of the principal and that of the agent; that is:

$$\max_{a(\delta), w(y,\delta)} [\delta v(w(y, \delta), a(\delta)) + (1 - \delta) u(y, w(y, \delta))],$$

subject to

$$\int_Y v(w(y, \delta), a(\delta)) f(y; a) dy \geq \int_Y v(w(y, \delta), a'(\delta)) f(y; a'(\delta)) dy \quad \forall a'(\delta) \in A,$$

$$0 \leq w(y, \delta) \leq y \quad \forall y \in Y.$$

A.3.3 The Multi-Objective Dynamic Model With No Bargaining Power Dynamics (MODPA1)

A third reference model is the multi-objective dynamic principal-agent model with no dynamics of bargaining power. In particular, we analyze the optimal contractual arrangements having the agent's bargaining power, δ, as the model's state variable, but without the implications of explicitly including a law of motion for δ. That is, in this version of the model we analyze the maximization of the Pareto Weights function of the expected discounted utility of the principal and that of the agent, subject to the feasibility and incentive compatible constraints. This problem is formulated as follows:

$$\max_{a(\delta),w^{\delta}(y^{\delta},W),\overline{V}^{\delta}(y^{\delta},W),\overline{U}^{\delta}(y^{\delta},W)} [\delta V^{\delta} + (1-\delta)U^{\delta}],$$

where:

$$U(\delta) = \int_{Y} [y - w(y,\delta) + \beta \overline{U}(y,\delta)] f(y; a(\delta)) dy,$$

$$V(\delta) = \int_{Y} [v(w(y,\delta), a(\delta)) + \beta \overline{V}(y,\delta)] f(y; a(\delta)) dy;$$

subject to

$$a(V) \in argmax_{a'} \quad \hat{V}(a', V),$$

$$0 \leq w(y,\delta) \leq y \quad \forall y \in Y,$$

$$\delta \in (0,1),$$

$$(\overline{U}(y,\delta), \overline{V}(y,\delta)) \in \mathscr{W}(\delta) \quad \forall y \in Y.$$

B Electronic appendix

Codes and further figures are provided in electronic form. Available at `https://github.com/genarobasulto/Tesis_GBM`.

References

Arnold, B. C. (2014). Pareto distribution, *Wiley StatsRef: Statistics Reference Online* pp. 1–10.

Bouteska, A. and Mefteh-Wali, S. (2021). The determinants of ceo compensation: new insights from united states, *Journal of Applied Accounting Research* .

Clementi, G., Cooley, T. and Di Giannatale, S. (2010). A theory of firm decline, *Review of Economic Dynamics* **13**: 861–885.

Di Giannatale, S., Quetzalcoatl, I. and Basulto, G. (2021). The dynamics of bargaining power in a principal-agent model.

Edmans, A. and Gabaix, X. (2009). Is ceo pay really inefficient? a survey of new optimal contracting theories, *European Financial Management* **15**(3): 486–496.

Edmans, A. and Gabaix, X. (2016). Executive compensation: A modern primer, *Journal of Economic Literature* **54**(4): 1232–87.

Ehrgott, M. (2005). *Multicriteria Optimization*, Springer.

Frydman, C. and Jenter, D. (2010). Ceo compensation, *Annual Review of Financial Economics* **2**(1): 75–102.
URL: *https://doi.org/10.1146/annurev-financial-120209-133958*

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

Gayle, G. L. and Miller, R. A. (2009). Has moral hazard become a more important factor in managerial compensation?, *American Economic Review* **99**(5): 1740–69.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Publishing Company.

Grossman, S. and Hart, O. (1983). An analysis of the principal-agent problem, *Econometrica* **51**: 7–45.

Hernández-Lerma, O. and Romera, R. (2004). The scalarization approach to multiobjective markov control problems: Why does it work?, *Applied Mathematics and Optimization* **50**(3): 279–293.

Hölmstrom, B. (1979). Moral hazard and observability, *The Bell Journal of Economics* pp. 74–91.

Hölmstrom, B. (2004). Pay without performance and the managerial power hypothesis: A comment, *J. Corp. L.* **30**: 703.

Mandelbrot, B. (1960). The pareto-levy law and the distribution of income, *International economic review* **1**(2): 79–106.

Murphy, K. J. and Sandino, T. (2010). Executive pay and "independent" compensation consultants, *Journal of Accounting and Economics* **49**(3): 247–262.

Pareto, V. (1964). *Cours d'économie politique*, Vol. 1, Librairie Droz.

Pickands III, J. (1975). Statistical inference using extreme order statistics, *the Annals of Statistics* pp. 119–131.

Rogerson, W. (1985). The first-order approach to principal-agent problems, *Econometrica: Journal of the Econometric Society* pp. 1357–1367.

Spear, S. and Srivastava, S. (1987). On repeated moral hazard with discounting, *The Review of Economic Studies* **54**(4): 599–617.

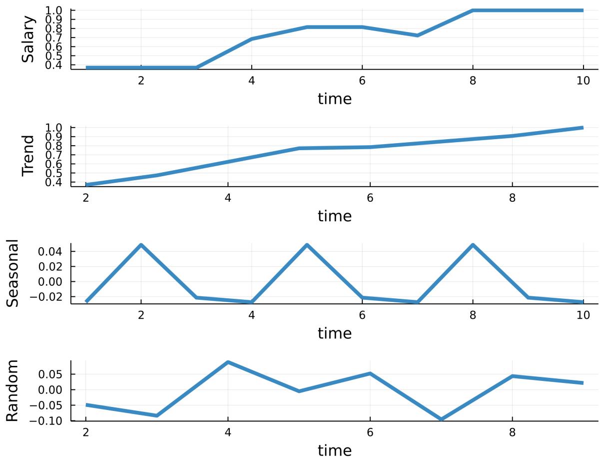
Sur, S., Magnan, M. and Cordeiro, J. (2015). Disentangling ceo compensation: A simultaneous examination of time, industry, and firm-level effects, *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration* **32**(1): 30–46.

Tervio, M. (2008). The difference that ceos make: An assignment model approach, *American Economic Review* **98**(3): 642–68.

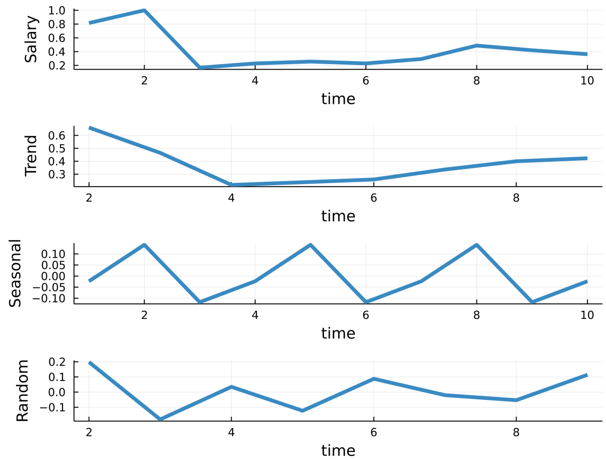
Wang, C. (1997). Incentives, ceo compensation, and shareholder wealth in a dynamic agency model, *Journal of Economic Theory* **76**(1): 72–105.

Zipf, G. K. (1949). Human behaviour and the principle of least effort.

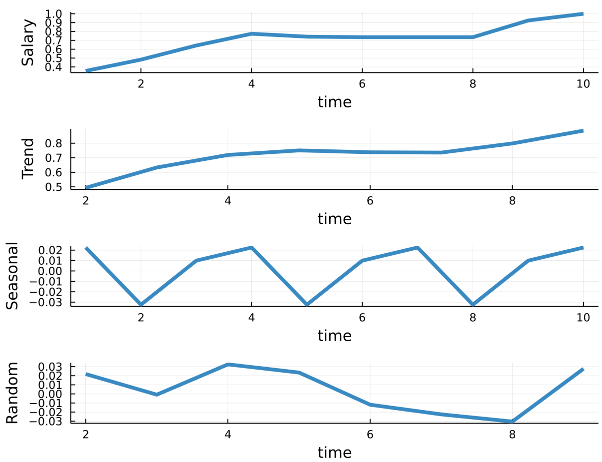
Figure 11: Decomposition of Observed Salary Time Series.



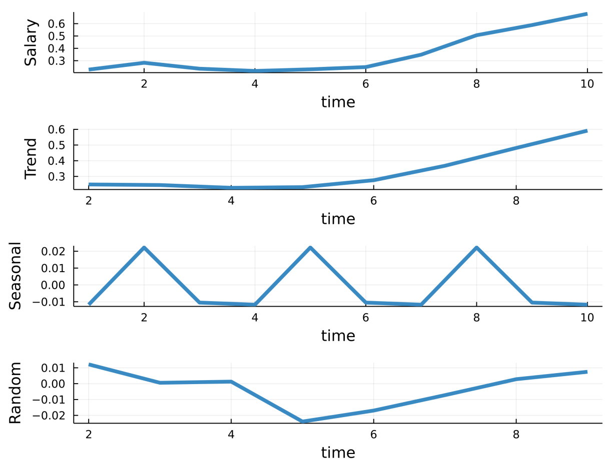
(a) BancFirst Corporation.



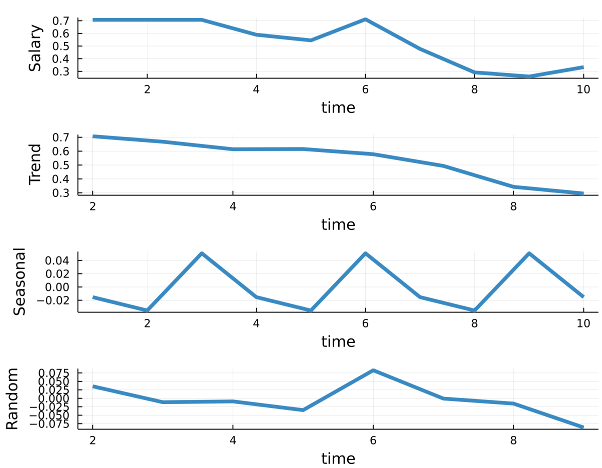
(b) Qurate Retail, Inc.



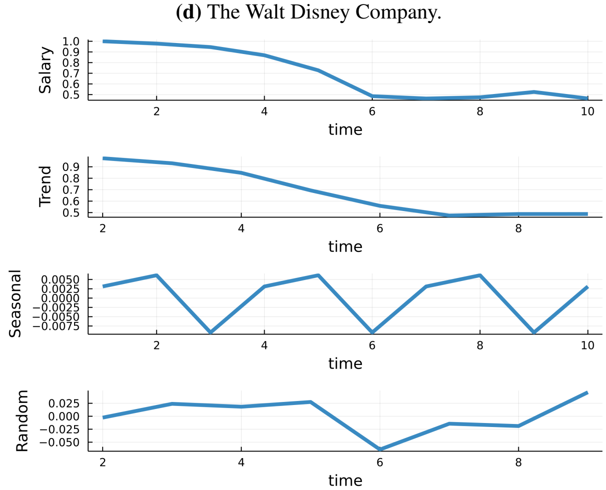
(c) ArcelorMittal.



(d) The Walt Disney Company.



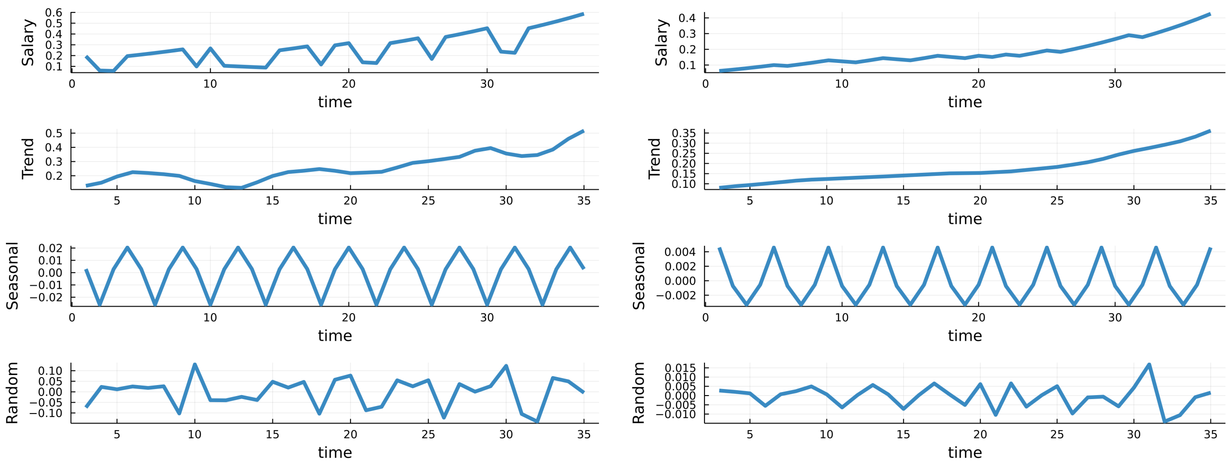
(e) Hornbeck Offshore Services, Inc.



(f) The AES Corporation.

Source: Own Elaboration with data from S&P.

Figure 12: Decomposition of Simulated Salary Time Series.



(a) Pareto Distribution. **(b)** Conventional Distribution.

Source: Own Elaboration with data obtained via numeric simulation.