Organizational Capital, Managerial Heterogeneity, and Firm Dynamics.

Wouter Dessein and Andrea Prat
Columbia University*

August 7, 2017

Abstract

We argue that economists have studied the role of management from three perspectives: contingency theory (CT), an organization-centric empirical approach (OC), and a leader-centric empirical approach (LC). To reconcile these three perspectives, we augment a standard dynamic firm model with *organizational capital*, an intangible, slow-moving, productive asset that can only be produced with the direct input of the firm's leadership and that is subject to an agency problem. We characterize the steady state of an economy with imperfect governance, and show that it rationalizes key findings of CT, OC, and LC, as well as generating a number of new predictions on performance, management practices, CEO behavior, CEO compensation, and governance.

*We thank Jacques Cremer, Luis Garicano, Hugo Hopenhayn, Navin Kartik, Qingmin Liu, Michael Raith, Jorgen Weibull, Pierre Yared, and participants to seminars at University of Athens, Columbia, Cowles Foundation Summer Conference, EIEF, EUI, LSE, Rochester, St Gallen, and Toulouse for useful suggestions. We are particularly grateful to Sui Sun Cheng for his suggestions on how to analyze the recurrence equation.

1 Introduction

A number of empirical studies, exploiting different data sets, employing different methodologies, and covering different countries have found sizeable and persistent performance differences between firms that operate in the same industry and use similar observable input factors (Syverson 2011). For instance, within narrowly specified US manufacturing industries, establishments at the 90th percentile make almost twice as much output with the same input (Syverson 2004).

One possible explanation for this puzzling observation is that the variation in outcomes is due to a variation in management (Gibbons and Henderson 2013). In turn, management comprises both the management practices that firms put in place and the managerial human capital that they employ. This paper is concerned with the question: Where do differences in management practices and managerial capital come from?

Economists have approached this question from three different angles. The first approach, which we shall refer to as *contingency theory (CT)*, is a natural extension of production theory. Both managerial practices and managerial human capital are production factors and the firm should select them optimally given the business environment it faces. Lucas (1978) is the seminal application of CT to managerial human capital. There is a market for managers where supply is given by an exogenous distribution of managers of different talent and demand is given by an endogenous distribution of firms In equilibrium, the more talented managers are employed by the firms that need them more. This model can be used to explain the allocation of CEOs to companies according to firm size (Tervio (2008) and Gabaix and Landier (2008)).[1] CT encompasses both managerial talent and management practices, and it can take into account synergies with other productive factors: Milgrom and Roberts' (1995) theory of complementarity in organizations develops general techniques to model these synergies. In sum, CT yields two powerful testable predictions: (i) If the solution to the production problem is unique, similar firms should adopt similar management practices and should hire similar managerial talent; (ii) If the production problem has multiple solutions, similar firms may adopt different management practices and/or hire different managers, but this variation will not correlate with their profitability.

While CT has an explicit theoretical foundation, the other two approaches are mainly

[1] See also Garicano and Rossi-Hansberg (2006), where more talented managers are matched with more talented subordinates, resulting in more productive firms.(ignoring wages).

empirical. We will refer to the second one as the *organization-centric empirical approach (OC)*. Ichniowski et al (1997) pioneered this approach in economics. They undertook a detailed investigation of 17 firms in a narrowly defined industry with homogeneous technology (steel finishing) and documented how lines that employed innovative human resource management practices, like performance pay, team incentives, and flexible assignments, achieved significantly higher performance than lines that did not employ such practices. Bloom and van Reenen (2007) developed a survey tool to measure managerial practices along multiple dimensions. Their paper and subsequent work have documented both a large a variation in management practices across firms within the same industry and the ability of that variation to explain differences between firms on various performance measures, including profitability.[2] These results are robust to the inclusion of firm-level fixed effects (Bloom et al 2016) and they survive the inclusion of detailed employee-level information (Bender et al 2016). In sum, OC has shown that similar firms adopt different management practices and that this difference matters for performance. If one wishes to reconcile these findings with CT, one should argue that those seemingly similar firms actually have different unobservable costs or benefits of adopting "better" practices. If one instead uses these findings to argue that CT fails in a systematic way, it would be useful to understand why so many firms in so many industries and geographies do not adopt profit-maximizing practices.

The third methodology is the *leadership-centric empirical approach (LC)*. In a nutshell, some firms are better run and they perform better because they have better CEOs. A growing literature, employing different data sets and different methodologies, show that the identity of the CEO can account for a significant portion of firm performance (Bertrand 2009). Among others, Bertrand and Schoar (2003) identify a CEO fixed effect, Bennedsen et al (2007) show that family CEOs have a negative causal effect on firm performance, Kaplan et al (2012) document how CEOs differ on psychological traits and how those differences explain the performance of the firms they manage, Bandiera et al (2016) perform a similar exercise on CEO behavior and show it accounts for up to 30% of performance differences between similar firms. Thus, LC can be seen as the parallel of OC, applied to top management rather than practices, which raises the same question: Can we reconcile the LC facts with CT?

[2] Graham et al. (2016), based on surveys of over 1300 CEOs and CFOs of US companies, obtain similar results with respect to the heterogeneity and effectiveness of corporate culture. We view both management practices and corporate culture as being part of a firm's organizational capital.

Leaving aside CT, it is also natural to ask if there is a link between OC and LC. Do CEO characteristics matter in the adoption of management practices? Are firms with certain management practices more likely to hire a certain type of CEO?[3]

This paper is an attempt to reconcile these three approaches in one theoretical framework. The objective is not develop a general, realistic model of management and managers, but rather to show that some of the essential lessons from CT, OC, and LC can be distilled in a set-up that requires only a small number of deviations from a standard dynamic firm model.

The premise of the paper is that the performance of a firm depends on its *organizational capital.* This concept is meant to encompass any intangible firm asset with four properties: (i) It affects firm performance; (ii) It changes slowly over time; (iii) Being intangible, it is not perfectly observable; (iv) It must be produced at least partly inside the firm with the active participation of the firm's top management.

We define organizational capital in a broad way, so it can conceivably include relational contracts, corporate culture, as well as other intangible, slow-moving characteristics that affect productivity. In particular, it may capture components of the management practices analyzed by Bloom et al (2016), which arguably affect firm performance (i) and are slow-moving (ii). In support of the imperfect observability condition in (iii), note that the correlation rate of two independent and almost simultaneous measurements of management scores within the same plant is 45.4% (Bloom et al 2016).

Instead, condition (iv) is mostly novel to economists. It is a tenet of an influential stream of management literature that includes Drucker (1967) and Kotter (2001). It is best summarized by Schein (2010) when he argues that "leadership is the source of the beliefs and values" of employees, and shapes the organizational culture of the firm, which ultimately determines its success or failure." In that perspective, some firms end up with leaders who are more capable and/or willing to act in a way that increases the firm's organizational capital. Note that this view of leadership is much more precise than simply saying that some CEOs generate more profits than others for some unspecified reason. It identifies a particular mechanism, the growth of intangible organizational assets, through which long-term value creation occurs in ways that lead to a wealth of testable implications.

[3]In a sample of firms where both CEO behavior and management practices are measured, Bandiera et al (2016) find significant cross-sectional correlation between the two indices, and both have independent explanatory power on performance.

In the model we develop in this paper, organizational capital depreciates over time, but the CEO can devote her limited attention to increasing it. Alternatively, the CEO can spend her time boosting short term profit. The firm's profit-maximizing board hires a CEO in a competitive market for CEOs and can fire her at any time. Some CEOs are better than others at improving organizational capital. Firms are otherwise identical. They are born randomly and they die if their performance is below a certain threshold. There are no other factors of production or sources of randomness.

This barebone model is completed by three types of informational and contractual frictions. First, while cash flow can be measured almost continuously, the immaterial nature of organizational capital makes it harder to monitor. We assume that the board observes the cash flow stream immediately , but they only spot changes in organizational capital with a delay. Second, when a board hires a CEO they have limited information about the CEO's type, especially if the candidate has never held a CEO's position, namely they have an imperfect CEO screening technology. Third, the board is unable or unwilling to use high-powered incentives so low-type CEOs would quit voluntarily.

In equilibrium, firms would like to dismiss low-type CEOs but the latter hide their type for some time by boosting short-term behavior rather than investing in organizational capital. If the firm is lucky, it gets a good CEO who increases organizational capital and improves long-term performance (and retires at some point). If the firm is unlucky, it gets a bad CEO who depletes organizational capital and hurts long term performance before the firm fires her. This implies that the organizational capital of each firm follows a stochastic process punctuated by endogenous CEO transitions. The three sources of friction above are necessary and sufficient to generate this equilibrium.

The main technical result of the paper is the characterization of the steady state distribution of firms in this economy. At every moment, there coexist firms with different organizational capital, different leadership styles, and different performance, giving rise to stylized OC and LC cross-sectional patterns.

The main substantive result is a wealth of testable implications that bring together, in one model, some of the key patterns predicted or observed by CT, OC, and LC as well as new implications that bring together the three approaches. On the CT front, our model displays the performance heterogeneity and persistence predicted by Hopenhayn (1992) as well as a power law at the top of the distribution. In OC, our analytical results are consistent with the findings by Bloom and Van Reenen (2007) and others that (changes in) the quality of management practices are associated with (changes in) firm performance. This relationship is mediated by the quality of corporate governance. Regarding LC, we show that the CEO behavior, type, and tenure are all predictors of firm performance, as found in the CEO literature – with governance quality and the supply of managerial talent acting as a mediating variable.

Finally, the model predicts a wealth of new cross-sectional and dynamic interactions between CT, OC,and LC concepts: the tenure, behavior, type, and compensation of present and past firm's CEOs predict the current level and growth rate of the firm's organizational capital. As before, we perform comparative statics on governance quality and leadership supply. We also make predictions linking CEO career paths and the dynamics of organizational capital. For instance, a firm that was run in the recent past by a CEO who is currently employed by a larger firm should display an abnormally high growth in organizational capital and performance. Conversely, a firm whose last CEO was short tenured will have lower organizational capital and performance.

Our paper is structured as follows. The first part of the part microfounds a dynamic firm model. Section 2 introduces a continuous-time model of an infinitely-lived firm with organizational capital and endogenous CEO transitions. Section 3 characterizes the equilibrium of the model when frictions are sufficiently strong and shows that it gives rise to a stochastic process determining CEO behavior, CEO turn-over, organizational capital, and firm performance (Proposition 1).

Section 4 contains our main technical result: the characterization of the steady state equilibrium of a dynamic economy with a continuum of firms that behave according to the dynamic firm model of the previous section (Proposition 2). Given some assumptions about firm births and deaths, the equilibrium distribution of firms obeys a recurrence equation, whose steady state admits one closed-form solution. For sufficiently high performance levels, the solution satisfies an approximate power law.

Section 5 explores the testable implications of the steady state characterization and shows that it reconciles key findings of CT (heterogeneity and persistence of firm performance), OC (cross-sectional and longitudinal relationship between management practices and firm performance), and LC (relationship between CEO behavior/type and performance). The section also analyzes the role of corporate governace and presents novel testable implications linking OC and LC variables.

Section 6 introduces observable heterogeneity in CEO quality. Suppose CEOs can live for more than one period and work for more than one firms. The market for CEOs will then be segmented into untried CEOs, successful CEOs, and failed CEOs. In equilibrium failed CEOs are not re-hired, untried CEOs work for companies with low organizational capital, and successful CEOs receive a compensation premium to lead companies with high organizational capital. The extension leads to additional predictions: a panel regression run on data generated by this model would yield CEO fixed effect coefficients; however, because of the endogenous assignment of CEOs to companies, such coefficient would under-estimate the true effect of individual CEOs on firm performance. The section also yields novel predictions on the dynamic relationship between CEO compensation, CEO career, firm performance, and the growth of organizational capital. Section 7 briefly concludes.

1.1 Literature Review

In addition to the many papers mentioned above, our paper is related to a few others. At least since Hopenhayn (1992) and Erickson and Pakes (1995), economists have emphasized how firm specific sources of uncertainty can result in firm dynamics and long-term productivity differences between ex ante similar firms in the same industry. We follow Hopenhayn in analyzing the steady state outcome of this dynamic process, but we micro-found one of the possible sources of the idiosyncratic productivity shocks by introducing managerial skill heterogeneity and moral hazard in the building of a firm's organizational capital (management practices, culture,. . .). As such, we are able to link the distribution of firm productivity to corporate governance and the supply of managerial talent, and make predictions which directly link managerial talent with organizational capital and firm productivity.

One of the main objectives of this paper is to reconcile the OC, LC and CT perspectives discussed above. Within OC, Bloom et al. (2016) consider a dynamic model which attempts to reconcile CT with OC. In their model, firms make costly investments in a 'stock of management'.[4] One of the key result of the paper is that the empirical patterns they observe can be

[4] In Bloom et al. (2016), more management is always better. They refer to this perspective as 'Management

rationalized by assuming a heterogeneous initial draw of management quality: firms are born with a random level of management quality, and this continues with them throughout their lives. As in Lucas (1978), this initial variation is not explained within the model and – to fit the data – it must be of the same order of magnitude of the observed (endogenous) variation in management practices. We follow Bloom et al. (2016) in thinking of management quality – an example of organizational capital – as a slow-moving asset. However, we differ in that we fully endogenize this asset and in so doing we create a role for corporate leadership. This has two benefits: there is now a three-way link between CT, OC, and LC and the observed variation in organizational capital can now be explained entirely within the model without invoking exogenous differences between firms.

Within LC, Bandiera et al (2016) consider an assignment model where different types of firms are more productive if they are matched to CEOs that choose the right behavior for that firm. In the presence of limited screening and poor governance, some firms may end up with the wrong CEO, thus generating low performance. This paper uses a similar building block, but combines it with organizational capital in a dynamic firm model and studies steady-state properties.

On the theory front, there are a number of other models where similar firms end up on different performance paths. Chassang (2010) and Li, Matouschek and Powell (2017) show how performance differences between (ex ante) identical firms may arise because of path-dependence in (optimal) relational contracts. In Chassang, for example, differences in a firm's success in building efficient relational contracts determines productivity differences. In the same vein, success (or lack thereof) in hiring a good CEO determines whether a firm's organizational capital grows or deteriorates in our model. Path dependence in developing efficient rules for employee behavior also result in performance differences in Ellison and Holden (2014). The above papers, however, do not derive a steady state distribution of firm performance, nor do they show how differences in corporate governance and the supply of managerial talent result in differences in managerial practices and organizational capital.[5]

as a Technology' and contrast this with 'Management as a Design', a setting in which there are no good or bad management practices.

[5]Halac and Prat (2016) model one type of organizational capital: a costly but imperfectly observable monitoring technology that must be maintained by top management. In equilibrium, identical firms may end up on different performance paths.

Board et al. (2016) propose a model where firms with higher levels of human capital are better at screening new talent, creating a positive feedback loop. Similarly, in Powell (2016), firms which earn higher competitive rents have the credibility to adhere to more efficient relational contacts with their employees, creating a positive feedback loop. While both models offer powerful theories as to why imitation is difficult and why some firms may have a sustained productivity advantage over others, ex ante identical firms obtain identical payoffs. Indeed, in Board et al. (2016), identical firms may choose different dynamic paths for human capital accumulation and productivity, but pay a price for it which makes them indifferent. Powell (2016) requires initial differences between firms which are, however, magnified through his mechanism.

Finally, our paper is related to a literature in corporate finance on managerial short-termism (Stein (1989)). Most of this literature is focused on how different financial contracts (e.g. short-term versus long-term debt) trade-off a desire for early termination of unprofitable projects with the need to provide adequate incentives for long-term investments (Von Thadden (1995)). In contrast, we study the consequences of heterogeneity in managerial short-termism on the productivity dispersion of ex ante identical firms. Indeed, in our paper, bad managers are able to temporarily mimic the performance of good managers by boosting short-term performance at the expense of long-term investments in organizational capital. Our model further differs from classic models of managerial short-termism in that only bad managers engage in short-term behavior.

2 A Dynamic Model of Firm Performance

We propose a dynamic (continuous time) model of firm performance where profits at time t are a function of the firm's organizational capital Ω_t. This organizational capital includes the quality of the firm's management practices and management system, its culture and norms and so on. The firm has a CEO whose responsibility it is to maintain and grow this organizational capital, denoted as behavior $x = 1$, but can shirk on this responsibility and engage instead in activities which boost short term performance, denoted as behavior $x = 0$.[6]

In the long-term behavior ($x = 1$), the CEO might be building a management system

[6] One key simplifying assumption is that the CEO chooses her behavior once and for all at the beginning of her tenure. The assumption is discussed – together with other limitations of the model – after Proposition 1.

and provides supervision and motivation to workers. In the short-term behavior ($x = 0$), the CEO might instead spend her time boosting productivity immediately. For example, the CEO could be monitoring operations directly as opposed to creating an accountability system, or going on sales pitches as opposed to incentivizing/training sales managers. Central to our analysis is that there are two types of CEOs, good and bad, who differ in their managerial ability to build organizational capital.

Formally, the firm's performance or flow profit at time t is given by

$$\pi_t = (1 + b(1 - x)) \Omega_t, \tag{1}$$

where $b \in [0, \bar{b}]$ is a short-term boost to performance, as chosen by a CEO engaging in behavior $x = 0$. The firm's organizational capital is an asset that evolves according to:

$$\dot{\Omega}_t = (\theta x - \delta) \Omega_t,$$

where δ is the depreciation rate of managerial capital and $\theta \in \{\theta^L, \theta^H\}$ represents the CEO's managerial skill with $\theta^H > \theta^L$.

The model could easily be extended to include other production factors. For instance, one might have a standard formulation in which

$$\pi_t = (1 + b(1 - x)) \Omega_t f(K_t, L_t) - rK_t - wL_t - F, \tag{2}$$

where K_t is the amount of capital and r is its unitary cost, L_t is the amount of labor and w is its unitary cost, and F is a fixed cost. With this formulation, K_t and L_t would be chosen given the firm's organizational capital. Under standard assumptions, the optimal amount of capital and labor would be increasing in the value of the firm's organizational capital. The results presented in the rest of the paper would continue to hold, with minimal modfications. To keep notation to a minimum we abstract from other factors and use (1).

The owner (or board) maximizes long-term profits

$$\int_0^\infty e^{-\rho t} \pi_t dt$$

We assume that behavior 1 is optimal for both CEO types (θ^L large enough compared to $\bar{b}$). Hence, if the owner observed the CEO type she would always hire the high type and instruct her to choose $x = 1$.

The owner, however, does not observe the CEO type, the CEO's behavior $x \in \{0,1\}$, or the current level of the organizationcal capital immediately. They are observable with a delay R. The only variable the owner observes in real time is performance

The board appoints the CEO and she can fire him whenever she wants, but CEOs must retire after time T. The probability of selecting a high type θ^H is given by $p > 0$.

CEO's do not care about profits, but maximize tenure. When hired, the CEO chooses a management style and – for simplicity – we assume she cannot change it over time. We will discuss the relevance of this assumption in the next section.

3 CEO Behavior, CEO Turnover and Firm Performance

We first present the results of our simple model, which is based on a number of stark assumptions. At the end of the section, we discuss how robust the results are to modifications of the assumptions.

To gain intuition, suppose all CEOs behave naively. They all choose optimal behavior: $x = 1$. Managerial capital growth then equals

$$\dot{\Omega}_t = (\theta - \delta)\,\Omega_t,$$

and is thus faster for θ^H than for θ^L. As performance is given by $\pi_t = \Omega_t$, the performance growth rate is

$$\frac{\dot{\pi}_t}{\pi_t} = \theta - \delta$$

Note that in the latter case, the low type would immediately be spotted and fired. As we show next, this cannot be an equilibrium, as a low type CEO then has an incentive choose the short-term behavior.

Consider the case where good CEOs choose $x = 1$, but bad CEOs choose the short term behavior $x = 0$. While this causes organizational capital to depreciate, it allows the bad CEO to mimic the performance of good CEOs for a while. Normalizing t to 0 at the time of CEO hire, profits at time $t \in [0, T]$ are given by

$$\pi_t^H = \Omega_t^H = \Omega_0 e^{\left(\theta^H - \delta\right)t}$$

for the high type, whereas

$$\pi_t^L = (1 + b)\,\Omega_t^L = (1 + b)\,\Omega_0 e^{-\delta t}$$

for the bad type.

As long as $\left(1+\bar{b}\right)\Omega_t^L \geq \pi_t^H$, the bad type can mimic the good type by choosing a short-term boost $b \in \left[0, \bar{b}\right]$ so that $\pi_t^L = \pi_t^H$. Mimicking becomes unsustainable after a period:

$$\bar{t} = \frac{\ln\left(1+\bar{b}\right)}{\theta^H}.$$

Throughout the analysis, we assume that

$$T > \bar{t}. \tag{A1}$$

It follows that CEO type is identified for sure after $\bar{t}$ periods. That may come before or after the exogenous observational delay R. So, a bad CEO is fired after a period of $\bar{t} = \min\left(K, R\right)$. Good CEOs are kept until retirement $(T > \bar{t})$. Clearly, the above behavior is an equilibrium.

The following result holds:

Proposition 1 *A low-type CEO chooses behavior 0, is fired after a period* $\bar{t} = \min\left(K, R\right)$ *with* $K = \frac{\ln\left(1+\bar{b}\right)}{\theta^H}$*, and leaves a firm with a worse management system:*

$$\Omega_{\bar{t}}^L = \Omega_0 e^{-\delta\bar{t}} < \Omega_0.$$

A high-type CEO chooses behavior 1, serves until retirement, and leaves a firm with a better management system:

$$\Omega_T^H = \Omega_0 e^{\left(\theta^H - \delta\right)T}.$$

To illustrate the proposition, assume that $M_0 = 1$, $\theta^H = .10$, $\delta = .06$, $\rho = .05$, $\ln(1+\bar{b}) = .20$, $R = 3$, and $T = 5$. We therefore have that

$$\bar{t} = \frac{.20}{.10} = 2$$

so that a bad manager leaves after two years and leaves organizational capital that is $e^{-(.06)2} = 0.886$ times the capital she found. A good manager retires after 5 years and leaves an organizational capital that is $e^{(.04)5} = 1.221$ times what she found.

Figure 1 plots the organizational capital (plot a) and the performance (plot b) of a firm that hires a bad CEO, followed by another bad CEO, followed by a good CEO, followed by a bad CEO.[7]

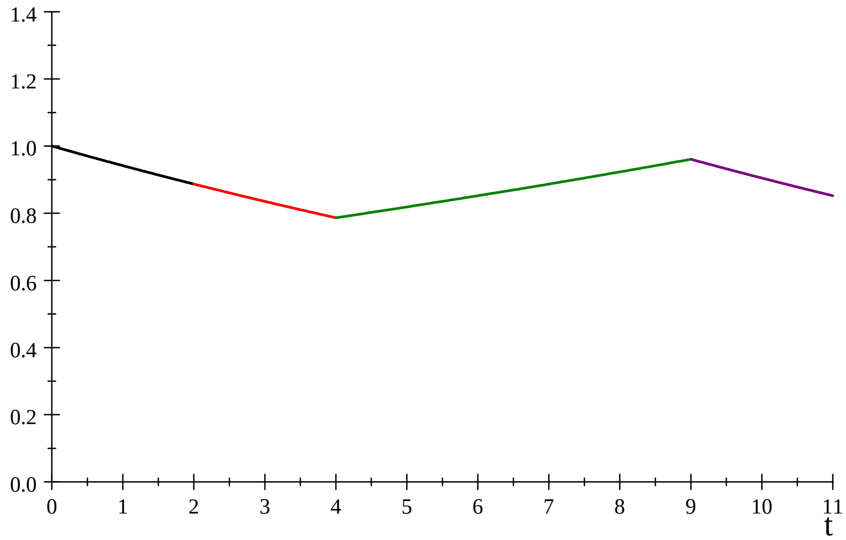

Figure 1: Organizational capital

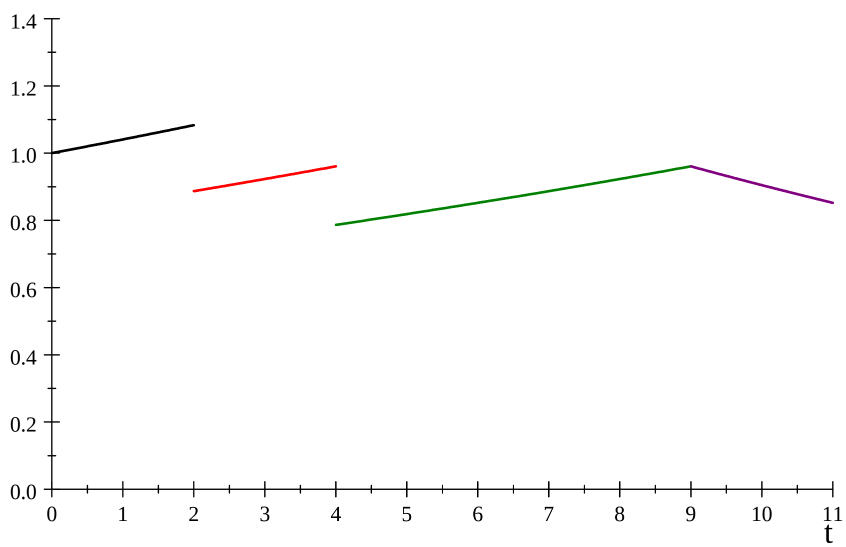

Figure 2: Evolution of performance

Proposition 1 depends on a number of stark assumptions we have made. As mentioned in the introduction, the results hinge on the presence of a serious agency problem within the company. In a frictionless environment, bad CEOs would either not be hired or leave immediately, in which case CEOs would only be high quality and there would be no leadership heterogeneity. Let us go over the various frictions we have assumed.

First, we posited that the owner is unable to screen CEOs based on their quality θ. If the owner had an effective screening technology, she would only hire the good ones. The extension of the model (Section 6) with various quality levels explores the possibility that CEOs can move from one firm to the other, in which case owners can learn something about the CEO's type from the performance of the firm they worked for previously.

Second, we assumed that the CEO receives a flat wage (normalized to zero). If the CEO's contract included a sufficiently strong performance-contingent component, a bad CEO could be incentivized to reveal his type right away.[8] This assumption can be assessed from a pragmatic perspective or a theoretical one. First and foremost, in practice it has been argued that, even in developed market economies such as the US, corporate governance is highly imperfect: the actual incentive schemes that CEOs receive are highly constrained and they do not align the CEO's interest with that of the firm (Bebchuk 2009). From a theoretical perspective, one can also show that enlarging the set of contracts available to the company may not weed out bad CEOs, because the incentive schemes that achieve this goal also increase the rent the firm must concede to all CEOs. This point is explored formally in Appendix II.

Third, we assumed that the owner does not observe organizational capital Ω_t directly. Obviously, if she does, she could kick out a bad CEO immediately. One could consider an alternative model where the owner observes a noisy continuous signal of organizational capital and will fire a CEO if enough evidence accumulates. The results would be qualitatively similar to the present model (but the analysis would be more complex – prohibitively so, at least for us, when we move to the aggregate level).

[7] Every time a bad CEO departs, the model predict a sharp drop in observed performance. This can be interpreted as a restatement of financial performance (some papers, like Hannes et al (2010) document the correlation between financial restatements and CEO turnover) or as an artifact of a model where performance is perfectly observable (see the discussion of assumptions below).

[8] For instance, suppose the CEO is offered a large stock option plan (a share of future profits): then, a bad CEO would rather resign right away in the hope that his replacement is of greater quality. It is possible to think about other schemes that would achieve the same result, like a golden parachute, backloaded compensation, etc.

Fourth, we assumed that the owner observes cash flow perfectly. This assumption too could be relaxed. As in the previous point, the resulting model would be much more complex. Having imperfectly observable performance would eliminate the stark negative effect on performance that we currently observe when a bad CEO departs.

Finally, we assumed the CEO cannot change her management style over time. This leads to equilibrium uniqueness. If the CEO were to be able to change her behavior over time, the equilibrium of Proposition 1 would still exist, but other perfect Bayesian equilibria may arise too. The good CEO could signal her type by first playing $x = 1$, then plays $x = 0$ before reverting back to $x = 1$. Since it would be sufficient for the good CEO to play $x = 0$ for an infinitesimal time, separation could occur (almost) immediately and a bad type would be fired (almost) instantly. Those immediate signaling equilibria would mainly be an artefact of our assumption that profits are perfectly observable and predictable. Unfortunately, adding noise to performance renders the analysis unwieldy very quickly. A more tractable way to eliminate signaling equilibria is to assume that the bad type is more productive at the short term behavior, that is $b \in \{b^L, b^H\}$ and the CEO's type is either (θ^L, b^H) or (θ^H, b^L).[9] Our assumption that CEOs needs to commit to a particular management style once hired achieves the same goal and keeps the model simple.

4 Steady-State Distribution of Firm Performance

Now that we have characterized the equilibrium behavior of an individual firm, we analyze aggregate behavior. We assume there is a continuum of firms and:

Assumption S1: *A firm dies whenever its performance goes below a certain profit level* π_0.

Assumption S2: *At each moment a measure* B *of new firms are born as spin-offs of existing firms. The spin-offs are clones of existing transitioning firms (firms who are changing their CEOs) and they inherit the organizational capital level of the firm they originate from.*

Neither S1 or S2 is necessary for the substance of our results, but their combination leads to a closed-form steady state. One could instead assume that there is an arbitrary birth

[9] Without loss of generality, one could also introduce a third type of manager (θ_L, b_L) which is lousy at both behaviors. It suffices then that both other types engage in signalling for a (infinitessimal) short time right after being hired, for such a type to be immediately discovered and fired.

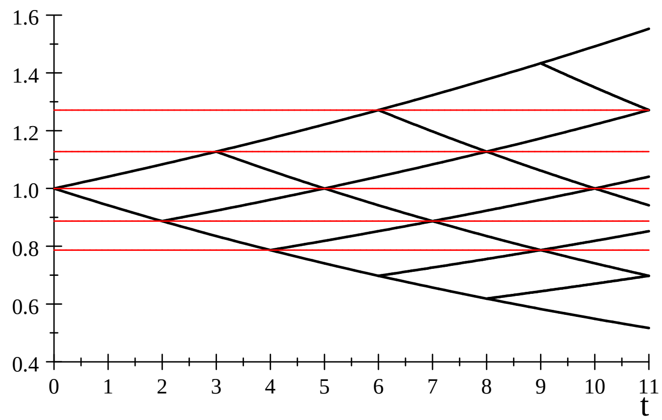

Figure 3: Possible organizational capital paths

function $\beta(\Omega)$ that determines the density of firms born at a given level and provide a numerical solution to the problem.

We first perform analysis under the following assumption:

Assumption S3: *The effect of a bad CEO exactly undoes the effect of a good CEO*:

$$\begin{aligned} \Omega_t e^{(\theta^H - \delta)T} e^{-\delta \bar{t}} &= \Omega_t \\ \iff &\quad (\theta^H - \delta)T = \delta \bar{t} \end{aligned}$$

Assumption S3 combined with Proposition 1 implies that all firms will experience transitions at a stable countable number of performance levels. This greatly simplifies the exposition of the results. The extension of the findings to cases beyond S3 involves a time-dependent rescaling of performance, which will become easier to present, once the baseline case is understood.

Figure 3 illustrates possible organizational capital paths when $\Omega_0 = 1$. Thanks to S3, all CEO transitions occur at a countable number of time-invariant levels.

At any CEO transition, performance and organizational capital are equal and fully known to the firm. The countable set of organizational capital levels at transition is thus the same as the countable set of performance levels at transitions, which we now define formally. Starting from the lowest performance level π_0, construct a set of performance levels Π as follows:

$$\Pi = \left\{ \pi : \exists l \in \mathbb{N} \text{ such that } \pi = \pi_k \equiv \pi_0 (1 + \Delta)^l \right\}$$

where Δ is the percentage improvement in organizational capital following a good CEO

$$1 + \Delta \equiv e^{\left(\theta^H - \delta\right)T}.$$

By Proposition 1, every firm that is born at a performance level in Π will only experience CEO transitions at performance levels within Π.

4.1 Steady State Distribution of Firm Performance at CEO Transitions.

Our aim is to characterize the steady state distribution of firms by first characterizing the (discrete) steady state distribution of firms over performance levels $\pi_l \in \Pi$. We proceed as follows.

Consider a distribution of firms at a calendar time t. Each firm i can be characterized by two numbers: its organizational capital level Ω_t^i and the time elapsed since the last CEO transition τ_t^i, which we refer to as "transition age". A firm with transition age $\tau_t^i = 0$ is changing its CEO at time t. Given Proposition 1, those two information items uniquely determine the quality and behavior of the CEO and the firm performance, so we will ignore those variables in what follows.

Let $\phi_t(\Omega, \tau)$ denote the density of firms with organizational capital Ω and transition age τ. A steady state is a situation where the density ϕ_t is constant over time. A necessary condition for ϕ_t to be a steady state is that $\phi_t(\Omega, 0)$ is constant over time; in other words, the distribution of transitioning firms must be in steady state.

We already noted that for transitioning firms, $\Omega_t = \pi_t$ and π_t can only take the countable number of values in set Π indexed by $l \in \mathbb{N}$. The distribution of transitioning firms at t can therefore be described by a probability function over a countable set of performance levels. From now on, we focus exclusively on even numbered performance levels belonging to Π. For simplicity, we renumber $2l \to k$, and denote

$$\Pi^+ = \left\{\pi : \exists k \in \mathbb{N} \text{ such that } \pi = \pi_k \equiv \pi_0(1+\Delta)^{2k}\right\}$$

We begin by characterizing the steady state distribution of firms with even CEO transitions, denoted by

$$\tilde{\phi}_t(k) = \phi_t(\pi_k, 0) \, / \sum\nolimits_{\pi \in \Pi^+} \phi_t(\pi, 0)$$

Once that is done, it will be easy to extend the result to all firms (transitioning and non-transitioning).

4.1.1 Waves and periods

We start with some preliminaries. Let a *performance-time pair* (k, t) represent a possible performance level $\pi_k \in \Pi^+$ at calendar time t. We define *waves* and *periods* as follows:

Definition 1 (waves and periods) *The set of all possible performance-time pairs* $(k, t) \in \mathbb{N} \times \mathbb{R}$ *can be partitioned in a continuum of waves, indexed by* $r \in [0, 1)$:

$$w_r = \{(k, t) : \exists n \in \mathbb{Z} \text{ such that } t = \tau(r, n, k)\}$$

each consisting of a countable number of periods, indexed by $n \in \mathbb{Z}$:

$$w_{r,n} = \{(k, t) : t = \tau(r, n, k)\} \tag{3}$$

where $\tau(r, n, k) \in \mathbb{R}$ *is the calendar time that performance* k *is reached in period* n *of wave* r :

$$\tau : \quad (r, n, k) \quad \rightarrow \quad t = (r + n)(T + \bar{t}) + k(T - \bar{t})$$

For a given performance level $k \in \mathbb{N}$ and transition time $t \in \mathbb{R}$, a firm's wave $r \in [0, 1)$ and period $n \in \mathbb{Z}$ is uniquely determined by the equality $t = \tau(r, n, k)$. Hence, each performance-time pairs (k, t) belongs to one particular wave and period.

As the next lemma shows, given definition 1, a firm moves to the next period of the same wave at each even CEO transition. Hence, a firm remains associated with the same wave $r \in [0, 1)$ as long as it survives:

Corollary 1 *Let* $(k, t) \in w_r$ *and* $(k', t') \in w_{r'}$ *be two performance-time pairs associated with the same firm, then* $r = r'$. *Moreover, if* (k, t) *belongs to period* n *and* (k', t') *belongs to period* $(n + 1)$ *then* t' *occurs two CEO transitions after time* t.

Proof. Consider a firm which transitions to performance k at time t, and let (k, t) belong to period n of wave r. Two CEO transitions after time t, this firm will either "move down" to performance level $k - 1$ at time $t + 2\bar{t}$, "stay" at performance level $k' = k$ at time $t + T + \bar{t}$, or "move up" to performance level $k + 1$ at time $t + 2T$. Let (k', t') denote the time-performance pair of this firm two CEO transitions after it was at (k, t). In all three cases we have that $t' = t + T + \bar{t} + (k' - k)(T - \bar{t})$. It follows that (k', t') belongs to period $n + 1$ of the same wave as (k, t). ■

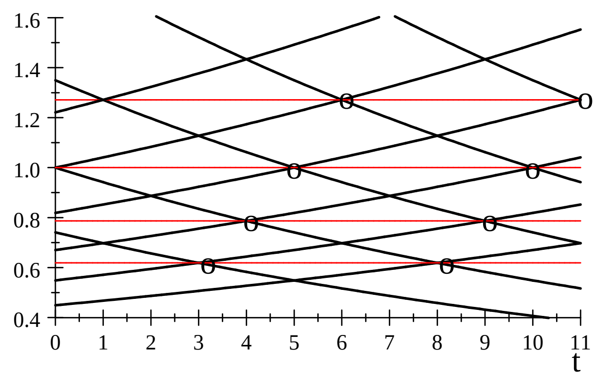

Figure 4: The circles in the figure are pairs (π_k, t) which belong to the same wave r and two consecutives periods n and $n' = n + 1$.

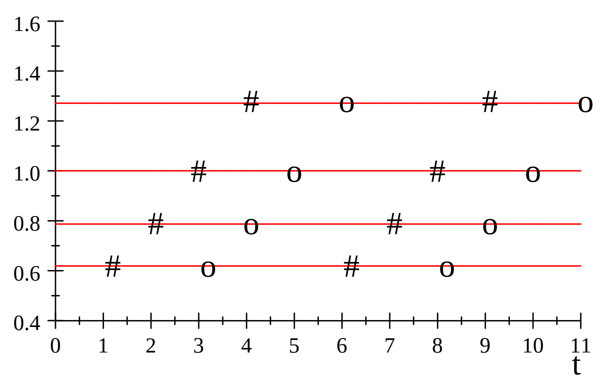

Figure 5: The #'s and o's in depicts (π_k, t) pairs of firms belonging to two different waves r and r'. Note that firms belonging to different waves transition to the same performance level at different times.

Given Lemma 1, we can say that each firm i belongs to (or is associated with) a particular wave $r \in [0, 1)$. Figure 4 illustrates one particular wave of firms. The circles in the figure are pairs (π_k, t) that belong to two consecutive periods of the same wave, say n and $n' = n + 1$. Figure 5 depicts (π_k, t) pairs of firms belonging to two subsequent periods of two different waves, say r and r'.

4.1.2 Recurrence Relation

Let $f_{r,n}(k)$ denote the measure of firms that transition to organizational capital level k in period n of wave r. Equivalently, $f_{r,n}(k)$ is the measure of firms with performance $\pi = \pi_k$ at time $t = \tau(r, n, k)$.

We are interested in characterizing the steady state of our economy. This is a situation where the distribution of firms is the same across period n and $n + 1$ and waves r and r':

$$f_{r,n}(k) = f_{r,n'}(k) = f_r(k) = f_{r'}(k) = f(k)$$

But this means that we can find the object we are interested in – the steady state distribution $\tilde{\phi}_t(k) = \tilde{\phi}(k)$ of firms transitioning at a particular calendar time t – by analyzing the steady state distribution of firms over a given wave r. Indeed, let $M = \sum_k f(k)$, then we must have that

$$\tilde{\phi}(k) = f(k)/M$$

One way to represent the steady state distribution of firms over a given wave is to modify figure 4 so the x-axis is expressed in wave periods rather than calendar time. We then obtain figure 6 below.

Note that, as depicted in Figure 6, a firm with performance level 1 in period 1 will have (i) performance level 2 in period 2 with probability p^2 (probability of two good CEOs), (ii) performance level 1 in period 2 with probability $2(1-p)p$ (probability of a good and bad CEO, or a bad and good CEO), (iii) performance level 0 in period 2 with probability $(1-p)^2$ (two bad CEOs).

We now proceed to characterize the steady state distribution $f(k) = f_{r,n}(k)$ of firms belonging to a given wave. Since firms can move up or down at most one performance level each period, we can write $f_{r,n+1}(k)$ as a function of $f_{r,n}(k)$, $f_{r,n}(k-1)$ and $f_{r,n}(k+1)$. To fix ideas, consider first the case in which no new firms are created, that is $B = 0$. For $k > 0$, we

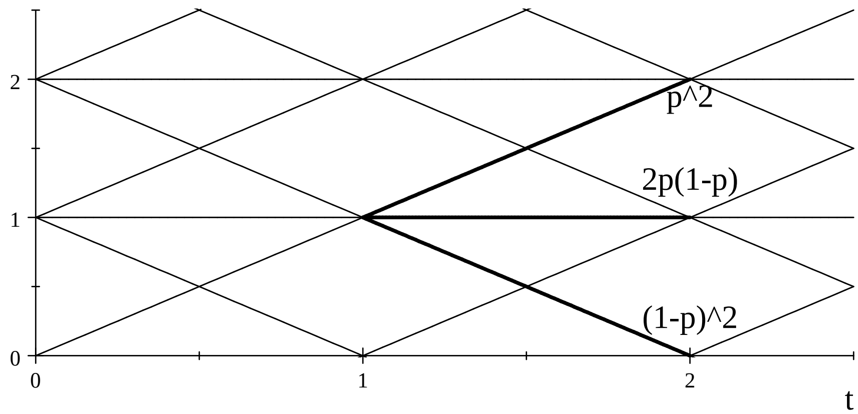

Figure 6: A conversion of Figure 4 in which: (1) The x-axis now represents wave periods rather than calendar time; (2) The y-axis represents the index k of set rather than performance value π_k. The legends corresponding to the bold segments indicate transition probabilities.

then obtain

$$f_{r,n+1}(k) = p^2 f_{r,n}(k-1) + 2p(1-p) f_{r,n}(k) + (1-p)^2 f_{r,n}(k+1)$$

Assume now that $B > 0$. Given Assumption S2, firms which transition at time t create a spin-off with the same organizational capital with probability B/M_t, where B is the constant mass of new firms that are born and M_t denotes the total mass of firms which transition at calendar time t. Taking into account such spin-offs, our transition function becomes, for $k = 1, 2, 3...$

$$f_{r,n+1}(k) = \left(1 + B/M_{\tau(r,n+1,k)}\right)\left[p^2 f_{r,n}(k-1) + 2p(1-p) f_{r,n}(k) + (1-p)^2 f_{r,n}(k+1)\right], \quad (4)$$

where $\tau(r, n+1, k)$ is the calendar time at which firms in period $n+1$ of wave r transition to level k.

Leaving aside M_τ, equation (4) can be seen a recurrence relation which is linear in two discrete variables, n and k (the discrete equivalent of a linear partial differential equation with two variables). However, of course, the presence of M_τ complicates things.[10]

If $f_{r,n}(k)$ is in steady state $f(k)$, then the total mass of transitioning firms in steady

[10] Indeed, M_τ is an endogenous variable and is composed of firms that belong to different waves as well as possible firms that belong to the same wave but in different periods (and with different performance levels).

state is constant too. Let us denote that level as M and

$$\gamma = B/M$$

In a steady state, (4) must hold for every wave. We can therefore drop r and re-write (4) as

$$f_{n+1}(k) = (1+\gamma)\left[p^2 f_n(k-1) + 2p(1-p)f_n(k) + (1-p)^2 f_n(k+1)\right], \tag{5}$$

This is a more tractable recurrence equation in two variables, n and k. There are two sets of boundary conditions:

$$\begin{aligned} f_n(0) &= 0 \text{ for every } n \\ f_n(1) &= \frac{B}{(1+\gamma)(1-p)^2} \text{ for every } n \end{aligned}$$

plus an initial distribution $f_0(\cdot)$. The first boundary condition says that firms die when they hit $k = 0$; the second one guarantees that exactly B firms die in every period (and hence in steady state B firms are born too).

Recurrence equations are sometimes used to represent heat diffusion processes in discrete time: a solid is subject to heating and cooling sources and we are interested in knowing the steady state temperature of different discrete points of the solid. One may wonder whether our problem corresponds to known diffusion problems. In this perspective, (5) can be loosely interpreted as a discrete version of the heat diffusion process of an imaginary one-dimension rod, with some additional features: (i) The rod begins at zero on the left side and it is unbounded on the right side; (ii) The diffusion parameter is asymmetric (as $p < \frac{1}{2}$, heat tends to flow left rather than right); (iii) The rod is heated along its length in a way that increases temperature at every point by rate γ per period; (iv) The left end of the rod is next to a powerful cooling source that keeps the temperature at zero. While this process is reminiscent of well-studied processes (especially in a continuous setting), we are not aware of any existing result that is directly applicable to the discrete case we are studying. We therefore proceed to characterize its steady state directly.

4.1.3 Steady State Analysis

If a steady state exists where $f_n(k) = f(k)$ for every k and every n, it must be that $f(\cdot)$ solves a second-degree difference equation in k:

$$f(k) = (1+\gamma)\left(p^2 f(k-1) + 2(1-p)\, p f(k) + (1-p)^2 f(k+1)\right). \tag{6}$$

It can be shown that (6) has a continuum of steady states, each associated with a different steady state spin-off rate $\gamma = B/M$. Unfortunately, (6) does not pin down γ which is an endogenous variable as it depends on the steady state mass of transitioning firms $M = \sum_k f(k)$. However, we will show that a simple refinement narrows the set of steady states down to one.

Note that in steady state, we must have that $\lim_{k\to\infty} f(k) = 0$.[11] Consider therefore the N-level version of our problem where we impose the boundary condition $f_{r,n}(k) = 0$ for $k > N$ with N a finite positive integer. In this finite version of our problem, organizational capital is bounded above by Ω_N. We say that *a steady state is reachable from below* if can be the limit of a sequence of steady states of the N-level version of our problem when $N \to \infty$. This requirement is natural because steady states that are not reachable from below require that at the beginning of time some firms already have unboundedly large organizational capital, which runs counter to our assumption that organizational capital is slowly accumulated with the help of the firm's leadership.

We prove:

Proposition 2 *In a steady state reachable from below, the mass of firms transitioning at performance level k is given by*

$$f^*(k) = M * \tfrac{(1-2p)^2}{(1-p)p} * k\left(\frac{p}{1-p}\right)^k$$

where M is the total mass of transitioning firms, given by

$$M = \frac{B}{\gamma^*} \equiv B\left(\frac{1-(1-2p)^2}{(1-2p)^2}\right)$$

The frequency of performance level k only depends on p :

$$\tilde{\phi}(k) = f^*(k)/M$$

Corollary 2 *The cumulative frequency distribution is given by*

$$\tilde{\Phi}(k) \equiv \sum\nolimits_{m\le k} \tilde{\phi}(m) = 1 - \left(\frac{p}{1-p}\right)^k \left(1 + k\frac{1-2p}{1-p}\right).$$

[11] If not, the mass of transitioning firms M is infinite, which cannot be a steady state.

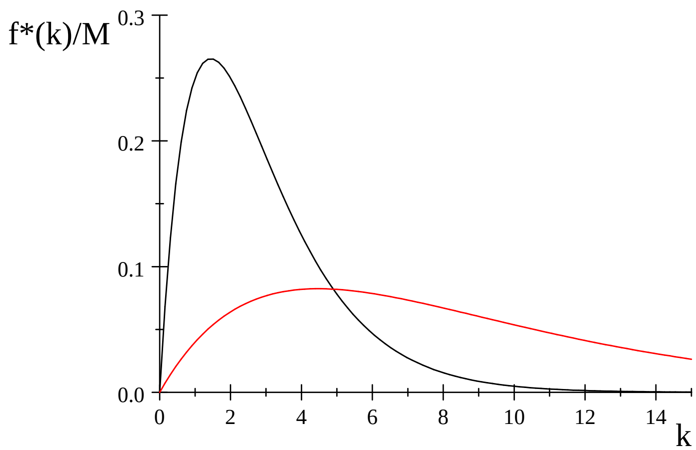

Figure 7: Frequency distribution $\tilde{\phi}(k) = f^*(k)/M$ (ignoring integer constraints) of transitioning firms, and this for $p = 1/3$ (black line) and $p = 4/9$ (red line). Recall that $\ln \Omega_k \sim k$.

An increase in p result in a right-ward shift of the performance distribution:

$$\frac{d}{dp}\tilde{\Phi}(k) < 0 \qquad \textit{for all } k \geq 1$$

Figure 7 plots the frequency distribution $\tilde{\phi}(k) = f^*(k)/M$ (ignoring integer constraints) of transitioning firms, and this for $p = 1/3$ and $p = 4/9$:

The proof of Proposition 2 (available in the Appendix) proceeds in a number of steps. We begin with the solution to the difference equation (6) together with the boundary conditions on the death threshold and the mass of births. As the difference equation takes the birth rate γ as exogenous, there is a distinct solution for every possible value of γ (the value of B just determines a rescaling of the whole distribution). We define a particular value of the birth rate:

$$\gamma^* \equiv \frac{(1-2p)^2}{1-(1-2p)^2}.$$

The proof then shows that for birth rates greater than γ^* the solution of the difference equation takes negative values for some positive integer k's. Intuitively, this is because a steady state cannot exist if the birth rate is too high because the distribution of firms would keep shifting rightward.

Finally, the proof shows that for any finite approximation of the original problem, the birth rate cannot be lower than γ^*. Formally, this is proven by deriving an upper bound to the eigenvalue of the transition matrix. If it were lower than that, then intuitively a steady state could not exist because the distribution would keep shifting left.[12] This proves that a necessary condition for a steady state reachable from below is that in equilibrium the birth rate is exactly γ^*. Once the value of γ^* is plugged into the general solution of (6), it yields the simple expression for $f^*(k)$ reported in Proposition (2).

To understand the steady state, consider the three forces that affect the size distribution of firms: a firm at size k can transition to $k+1$, k, or $k-1$ (and on average it drifts downward), low performers disappear when they hit the death threshold; a fixed mass of firms (not a percentage) is born at every transition time. The third force offsets the other two forces: if the total mass of firms became too low, the birth rate would go up. If the total mass of firms became too high, the birth rate would go down. This determines a unique steady state, where the outflow of firms through death equals the outflow of firms through birth and size distribution replicates itself over time.

One of the strongest empirical regularities on firm dynamics is that the right tail of the firm-size distribution follows a power law (Gabaix 2009, Luttner 2010). In line with this observation, Proposition 2 implies that the right tail of the distribution of organizational capital follows a power law. Abusing notation let us denote $\tilde{\phi}(\Omega_k) \equiv \tilde{\phi}(k)$. Since $\Omega_k = \Omega_0(1+\Delta)^{2k}$, we obtain that

$$\tilde{\phi}(\Omega_k) = \tfrac{(1-2p)^2}{(1-p)p} * \ln \Omega_k * \left(\frac{p}{1-p}\right)^{h * \ln \Omega_k}$$

where $h \equiv 1/(2\ln(1+\Delta))$. The following result holds:

Corollary 3 *In steady state, for $\Omega_k \in \Pi$ large, $\tilde{\phi}(\Omega_k)$ approximates a power law: There exists a $c > 0$ such that for Ω_k large $\tilde{\phi}(\Omega_k) \approx c * \Omega_k^{-\zeta}$ with $\zeta = h * \ln \frac{1-p}{p}$.*

Proof. See Appendix. ■

Figure 8 illustrates the convergence of the oganizational capital distribution to a power law for large levels of organizational capital. The power law approximation is a consequence of the underlying microfoundation, whereby performance follows a Markov chain.

Proposition 2 only applies to even CEO transitions. But it is easy to show that:

[12] There could exist steady states with lower birth rated but they would require a distribution of firms that is very skewed to the right to start with. Such a distribution cannot be reached from below.

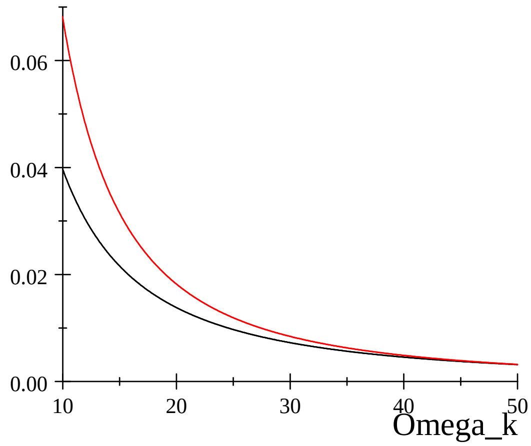

Figure 8: The black curve plots the frequency distribution $\tilde{\phi}(\Omega_k)$ for $\Omega_k \geq 10$ (parameters: $\Omega_0 = 1$, $\Delta = 0.2$ and $p = 1/3$). The red curve plots the asymptotic power- law distribution $c * \Omega_k^{-\zeta}$ with $\zeta \approx 1.9$.

Corollary 4 *Let $\hat{f}(k)$ be the steady state measure of firms with an uneven CEO transition at peformance level $\Omega = \Omega_k(1+\Delta)$. Then*

$$\hat{f}(k) = pf^*(k) + (1-p)f^*(k+1) \qquad \text{for } k = 0, 1, 2, ...$$

The steady state mass of firms with an uneven CEO transition is again given by M.

4.2 General Case: Good and Bad CEOs Have Different Absolute Effects.

The previous analysis was performed under Assumption S3, which states that the positive effect on organizational capital of a good CEO is exactly undone by the negative effect of a bad CEO, as depicted in Figure 9. We now remove this non-generic condition and allow the effect of a good CEO to be greater or smaller than that of a bad CEO. If, for instance, a good CEO has a larger absolute effect, then we have a situation as shown in Figure 10.

The red lines in Figure 10 can be called *neutral transition paths.* Assume without loss of generality that at $t = 0$, one of the neutral transition paths goes through $\Omega_0 = 1$. Then, that path is defined by

$$\pi_0(t) = e^{\frac{(\theta^H - \delta)T - \delta\bar{t}}{T + \bar{t}} t}$$

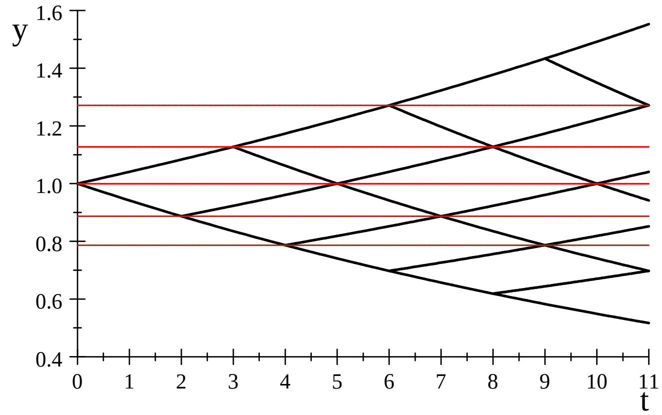

Figure 9: Possible performance paths when $(\theta^H - \delta)T = \delta\bar{t}$

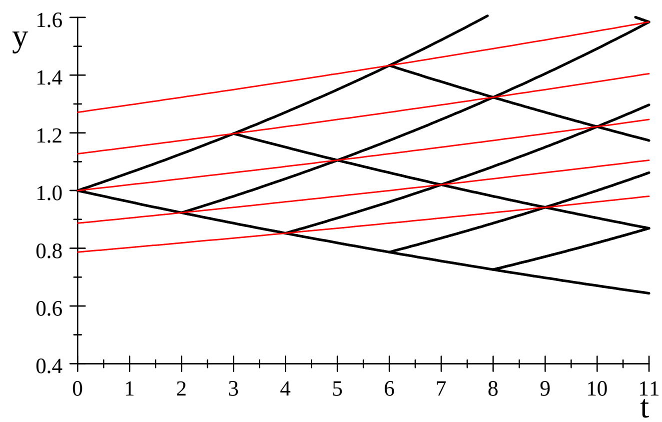

Figure 10: Possible performance paths when $(\theta^H - \delta)T > \delta\bar{t}$

and all other transition paths are defined by $\pi_l(t) = \pi_0(t)(1+\Delta')^l$ where l is an integer and

$$1+\Delta' = e^{(\theta^H-\delta)T - \frac{(\theta^H-\delta)T-\delta\bar{t}}{T+\bar{t}}} = e^{\theta^H \frac{\bar{t}T}{T+\bar{t}}}$$

All firms which experience a CEO transition at time t have a performance $\pi = \pi_l(t)$ for some $l \in \mathbb{Z}$. As before, we focus on even transition paths and relabel $k = 2l$. Consider therefore the set of (time-dependent) performance levels

$$\Pi^+(t) = \left\{\pi : \exists k \in \mathbb{N} \text{ such that } \pi = \pi_k(t) \equiv \pi_0(t)(1+\Delta')^{2k}\right\}$$

It is immediate to see the following

Proposition 3 *Suppose $\tilde{\phi}(k)$ is the steady state transition frequency of $\pi_k \in \Pi^+$ for an environment defined by $\left(B, \theta^H, \delta, T, \bar{t}\right)$ with $(\theta^H - \delta)T = \delta\bar{t}$. Then at time t, $\tilde{\phi}(k)$ is also the steady state transition frequency of $\pi_k(t) \in \Pi^+(t)$ for any environment defined by $\left(B', \theta'^H, \delta', T', \bar{t}'\right)$ where firms die whenever they reach $\pi_0(t)$ (and where θ'^H and δ' are consistent with the conditions in Proposition 1).*

Proof. Compute $\pi_k(t)$ and define level k as $\Omega_k = \pi_k(t)$. The recurrence equation is identical to that analyzed in Proposition 2. Proposition 3 applies to the steady state transition distribution over ordinal levels $k = 1, 2, 3....$ It also applies to time-variant cardinal levels defined by $\pi_k(t)$ with $k = 1, 2, 3....^*$ ■

The assumption in Proposition 3 that firms die whenever they reach transition path $\pi_0(t)$ can be replaced by:

Assumption S1': *A firm exits (or dies) whenever it is certain that no other firms in the economy have lower organizational capital.*

The steady state identified above is still a steady state under S1'. If all firms exit whenever their performance is smaller or equal than $\pi_0(t)$, then each firm which transitions to $\pi_0(t)$ at time t is known to be the firm with the lowest organizational capital in the economy – and hence exits by assumption S1'.

Note further that at the exact moment where firm i exits given Assumption S1', there is a discrete drop in the firm's expected organizational capital from $\Omega = p\pi_1(t) + (1-p)\pi_0(t)$ to $\Omega = \pi_0(t)$. An informal motivation for Assumption S1' is that such a drop in the perceived organizational capital of firm i results in a loss of confidence of customers, employees, financiers and so on, forcing firm i to exit.[13]

4.3 Steady State Distribution of Non-Transitioning Firms.

So far, we have only characterized the mass and frequency distribution of transitioning firms with peformance level $\Omega_k \in \Pi^+$. We now derive the cumulative distribution of all firms (transitioning and non-transitioning) with organizational capital $\Omega \leq \Omega_k$. For simplicity, we do it only for the special case where S3 holds.

We show first show that:

Lemma 1 *The steady state mass of all firms (transitioning and non-transitioning) is given by*

$$\bar{M} = 2M\left(pT + (1-p)\bar{t}\right)$$

Proof. From Proposition 2, there is a measure M of firms with an even CEO transition in steady state. Similarly, there is a measure M of firms with an uneven CEO transition. After a CEO transition firms are either led by a good CEO for time T or by a bad CEO for time $\bar{t}$. Hence, the steady state mass of firms is given by $\bar{M} = 2M\left(pT + (1-p)\bar{t}\right)$. ■

We want to characterize the cumulative distribution $\Phi(.)$ of this steady state mass of firms $\bar{M}$. As a reference, recall the cumulative frequency distribution $\tilde{\Phi}(.)$ of the steady state measure M of firms with an even CEO transition, $\tilde{\Phi}(k) \equiv \sum_{m \leq k} \tilde{\phi}(m)$, as derived in Proposition 2.

The next proposition shows that while $\tilde{\Phi}(k)$ only concerns firms with an even CEO transition, it is a good approximation for the probability that any firm's organizational capital Ω is smaller than Ω_k, where we include both transitioning and non-transitioning firms:

Proposition 4 *Let $\Phi(\Omega')$ be the cumulative distribution of all firms (transitioning and non-transitioning) with organizational capital Ω smaller or equal than Ω'. Then*

$$\tilde{\Phi}(k-1) < \Phi(\Omega_k) < \tilde{\Phi}(k)$$

with

$$\Phi(\Omega_k) = \tilde{\Phi}(k-1) + \tilde{\phi}(k)(1-p)\frac{\bar{t} + (1-p)\bar{t} + pT}{2((1-p)\bar{t} + pT)}$$

Proof. See Appendix ■

[13] Of course, uncountable other steady states are possible under S1.

5 Steady State Predictions

One goal of our simple model was to reconcile key predictions of the three existing approaches, CT, OC, and LC. This section lists the predictions that are consistent with each of the three perspectives. It also generates a number of new testable implications that cross over the three approaches.

The section is therefore divided into four sections: predictions consistent with CT, predictions consistent with OC, predictions consistent with LC, new predictions that cross over multiple approaches

5.1 CT Predictions

Hopenhayn (1992) and Erickson and Pakes (1995) posit that firms are subject to idiosyncratic shocks that affect their performance level. In steady state, we observe persistent performance differences (Gibbons and Henderson 2013). Namely: (i) A cross-section of otherwise identical firms exhibits different performance levels; (ii) The performance difference between any two firms is correlated over time.

Our model makes similar predictions. Let $\pi_{i,t}$ be the performance of firm i at time t. Based on Proposition 2, we immediately see

Proposition 5 *In steady state: (i) A cross-section of otherwise identical firms exhibits different performance levels ($Var(\pi_{i,t}) > 0$); (ii) The performance difference between any two firms is correlated over time (for any two firms i and j, and any $s > 0$, we have*

$$Corr(\pi_{i,t} - \pi_{j,t}, \pi_{i,t+s} - \pi_{j,t+s}) > 0$$

One of the strongest empirical regularities on firm dynamics is further that the right tail of the firm-size distribution follows a power law (Gabaix 2009, Luttner 2010). Building on Hopenhayn (1992) and Gabaix (1999), Luttner (2007) shows how – given the appropriate assumptions on the entry and exit process– models of firm dynamics with idiosyncratic shocks can generate such power laws. Similarly, as shown in Proposition 3, our model predicts that the right tail of the distribution of organizational capital follows a power law.

If one only observes firm performance and has no information over organizational capital or CEO variables, one would struggle to distinguish the model presented here from models like

Hopenhayn (1992) and Erickson and Pakes (1995) (except possibly for functional differences in the way (i) and (ii) manifest themseves). However, once organizational and managerial variables are observed, our model makes many more falsifiable prediction that we discuss in the next three subsections.

5.2 OC Predictions

Suppose now that the econometrician observes performance as well as organizational variables. The leading example is Bloom and Van Reenen (2007), where the form of organizational capital observed is the quality of management practices. They document how, after controlling for all observables, (changes in) the quality of management practices explains (changes in) firm performance. Moreover, the quality of management practices is correlated with corporate governance and the availability of managerial human capital.

These predictions are consistent with the relation between performance π and organizational capital Ω in our model. Again, take a steady state with a mass of otherwise identical firms. In our model, the quality of governance can be captured by $\bar{b}$, the ability of the CEO to create short term performance (the lower is $\bar{b}$, the faster bad CEOs are fired) and the availability of managerial human capital can be represented by θ^H. Recall that $\bar{t} = \frac{\ln(1+\bar{b})}{\theta^H}$

Finally define average performance growth, $E(\Delta\pi)$, as the average (instantaneous) growth of a randomly selected firm. From Proposition 2, we see:

Proposition 6 *In steady state:*

(i) In a cross-section of firms, performance and organizational capital are positively correlated: $Corr(\pi_{i,t}, \Omega_{i,t}) > 0$.

(ii) In a cross-section of firms, changes in performance are positively correlated with changes in organizational capital: For any $s > \bar{t}$,

$$Corr(\pi_{i,t+s} - \pi_{i,t}, \Omega_{i,t+s} - \Omega_{i,t}) > 0$$

(iii) Average performance growth is increasing in the quality of corporate governance and in the availability of managerial talent:

$$\frac{d}{db}E(\Delta\pi) < 0 \quad \textit{and} \quad \frac{d}{dp}E(\Delta\pi) > 0$$

Proof. (i) Immediate

(ii) The restriction $s > \bar{t}$ avoids a situation where performance growth is identical because bad and good CEOs cannot be distinguished in the initial period $\bar{t}$.

(iii) Recall that a firm gets a good CEO with probability p and grows at an instantaneous rate $\theta^H - \delta$ for T periods or gets a bad CEO with probability $1-p$ and grows at rate $-\delta$ for $\bar{t}$. So the average (instantaneous) growth of a randomly selected firm is

$$E(\Delta\pi) = \frac{p\left(\theta^H - \delta\right)T - (1-p)\,\delta\bar{t}}{pT + (1-p)\,\bar{t}}$$

where

$$\bar{t} = \frac{\ln\left(1+\bar{b}\right)}{\theta^H}.$$

Thus, it is easy to see that an increase in $\bar{b}$ produces the effects in (iii). ■

In steady state, ex ante identical firms have different levels of organizational capital, and this affects their performance. The heterogeneity is due to different leadership styles in the past. The same is true in terms of changes: firms whose last CEO was a good type experience a growth in both their organizational capital and their performance.

Note that in the model the effect of organizational capital on performance is causal. So, if an external intervention such as the one in Bloom et al (2013) were to increase $\Omega_{i,t}$, it would also increase performance $\pi_{i,t+s}$. Of course, the benefit of the model is that it explains where the heterogeneity in organizational capital comes from and it links it to another set of observables, as discussed below.

5.3 LC Predictions

The LC approach has studied the effect of CEO variables on firm performance (Bertrand and Schoar 2003, Bennedsen et al 2007, Kaplan et al 2012, Bandiera et al 2016). The CEO variables considered include the identity, the characteristics, and the behavior of the CEO.

The next section, where CEOs will be allowed to work at multiple firms and salaries are endogenous, will generate even more testable predictions on career trajectories and compensation patterns. However, for now, let us focus on the implications of the stylized model considered so far.

Consistent with the core of those observed patterns, our model predicts a connection between CEO variables and performance. In the equilibrium described in Proposition 1, good

CEOs behave differently, produce more organizational capital, generate better performance, and stay longer on the job. This in turn leads to a number of cross-sectional patterns:

Proposition 7 *(a) In steady state, firm i's current performance level $\pi_{i,t}$ is higher when past CEOs: (i) Chose the organization-building behavior rather than the short-term profit boost ($x_{i,t-s} = 1$ not 0); (ii) Were of the high type rather than the low type ($\theta_{i,t-s} = \theta_H$ not θ_L); (iii) Had longer on-the job tenure (T not $\bar{t}$).*

(b) In steady state, in a cross-section of firms, better governance (lower $\bar{b}$ or higher R) weakly increases the average behavior and type of the CEO, the tenure variance among CEOs, and average performance.

Note that predictions (a)(i) and (a)(ii) hold also in a probabilistic sense. If certain categories of CEOs are more likely to be high types and behave well, the firms run by those CEOs will in general have better performance and higher organizational capital. This rationalizes the findings of Bennedsen et al (2007) that family and professional CEOs impact long term performance differently.[14] Finding (a)(iii) is to the best of our knowledge untested but it is an immediate implication of a model where bad CEOs are more likely to be dismissed early. Finding (b) is consistent with the key findings of the literature on international differences in governance (Shleifer and Vishny 1997)

5.4 Predictions Linking OC and LC

As mentioned in the introduction, the OC and LC approaches have mostly operated in a separate manner. Our model suggests a number of testable implications involving OC variables and LC variables. CEOs play a part in growing or destroying organizational capital, which in turn determines performance. So our model predicts a lagged effect of CEO variables on organizational capital.

It is immediate to see that:

Proposition 8 *(a) In steady state, the rate of growth of organizational capital $\Omega_{i,t}$ is greater when the current CEO: (i) Chooses the organization-building behavior rather than the short-*

[14] Our model does not capture the feature that family CEOs are often difficult to dismiss. The result would hold a fortiori.

term proft boost ($x_{i,t} = 1$ not 0); (ii) Is of the high type rather than the low type ($\theta_{i,t} = \theta_H$ not θ_L); (iii) Has longer on-the job tenure (T not $\bar{t}$).

(b) Firm i's current organizational capital $\Omega_{i,t}$ is higher when past CEOs: (i) Chose the organization-building behavior rather than the short-term proft boost ($x_{i,t-s} = 1$ not 0); (ii) Were of the high type rather than the low type ($\theta_{i,t-s} = \theta_H$ not θ_L); (iii) Had longer on-the job tenure (T not $\bar{t}$).

(c) Controling for current organizational capital $\Omega_{i,t}$, past CEO variables have no predictive value on current firm performance π_{it}.

Organizational capital is a stock, while CEO behavior is a flow that influences the growth of the stock. Part (a) is an immediate consequence of this: organizational capital grows faster when at least one of the folllowing is true: the CEO behaves better, is a higher type, or has been there for longer (meaning that his type is more likely to be high). Part (b) is the cumulative correspondent of part (a): the current level of organizational capital is predicted by the type, behavior, and tenure of past CEOs. For instance, a firm that has experienced a sequence of short-lived CEOs is predicted to have a lower organizational capital.

Part (c) helps distinguish the present model from other stories that give the CEO a productive role. For instance, a charismatic CEO may have a direct motivating effect on employees that does not go through the growth of organizational capital. Such a (reasonable) model would create a direct link between CEO type/behavior and performance that would violate Part (c). Hence, Part (c) suggests a way of disaggregating the effect of the CEO between growing the organizational capital and affecting performance directly.

6 Model with Endogenous Wages and CEO Quality

So far we have assumed that CEOs only work once. What happens if a CEO can "prove herself" in one firm and then go to another firm? Which firms will hire better CEOs?

In this section, we first show a general result: if multiple CEO types are available, better CEOs will be hired by firms that already have more organizational capital.

We then apply this general result to a situation where CEOs can take a succession of jobs. In equilibrium, rookie CEOs are hired by low-performance firms. If they succeed, the move on to better firms. The salary differential between new and proven CEOs is determined in equilibrium.

6.1 The Marginal Value of CEO Quality

Reconsider our baseline model but assume that there are multiple categories of prospective CEOs. CEOs in category j have a p_j probability of being type θ_H and a $1 - p_j$ probability of being type θ_L. CEO compensation is endogenous. In equilibrium all CEOs in category j earn the same instantaneous wage w_j (we are maintaining the hypothesis that the only possible form of compensation is a constant per period wage).

We also assume that the cash flow boost b is not only bounded above by $\bar{b}$, but it does now allow the CEO to reach a performance level that is greater than that of high-quality CEO who chooses $x = 1$.[15] Without this additional assumption, a highly impatient firm might ask its CEO to engage in short term profit boosting. Alternatively, one could assume that b is not profit boosting, but covert borrowing: unbeknown to the board, the CEO borrows funds on behalf of the firm at instantaneous rate ρ that must be repaid by the firm when the CEO is fired. The result below holds a fortiori in the alternative scenario.

For the rest, the model is unchanged. We can show:

Proposition 9 *Consider a CEO with p' and a CEO with $p'' > p'$. Let Ω' and Ω'' denote the organizational capital levels of the firms employing the two CEOs respectively. If firms are sufficiently impatient (ρ is sufficiently high), in steady state $\Omega' \leq \Omega''$.*

Proof. See Appendix ■

Proposition 9 says that more promising CEOs must be hired by firms with higher organizational capital. A CEO with a higher p is more likely to protect the firm's organizational capital – something that is more useful when the size of the organizational capital is larger. The key assumption is that the effect of CEO behavior/type and organizational capital is multiplicative:

$$\dot{\Omega}_t = (\theta x - \delta)\,\Omega_t,$$

To reverse the effect, one must assume that

$$\dot{\Omega}_t = \theta x z\,(\Omega_t) - \delta \Omega_t,$$

[15] Formally, if t is the time when the CEO was hired and s is her tenure, $b_s \in \left[0, \min\left(\bar{b}, e^{(\theta^H - \delta)s}\right)\right]$.

where $z(\cdot)$ is a decreasing function. In that case, firms with lower organizational capital may hire more promising CEOs.

The requirement that ρ is sufficiently high is mainly technical and derives from the inability to characterize the value function of this problem.

Proposition 9 is related in spirit to results on assortative matching between CEOs and firm size (Tervio (2008) and Gabaix and Landier (2008)). There, more capable CEOs are matched with larger firms. Here, CEOs who are more likely to be good are matched with firms with a higher organizational capital. The connection would become direct if we used the production function in (2).

Of course, the fact that more expensive CEOs are more likely to be good types does not eliminate the stochastic element that underpins the our organizational capital process. Even expensive CEO may turn out to be bad and destroy organizational capital. The following section explores such dynamics.

6.2 Equilibrium with Proven CEO Quality

Consider now the endogenous allocation of CEO talent. As before, there are two types of CEOs, good and bad, and CEOs that are revealed to be bad can be fired at any time. We maintain S1-S3 above so that a bad CEO exactly undoes the effect of a good CEO on organizational capital. But rather than retiring after a period of time T, good CEOs may move to a different firm.[16] We assume that the type of a CEO is only partially persistent. A CEO with a low type always remains low. A CEO with a high type becomes a low type with probability κ at the end of a contract term T.

There are then three categories of CEOs. A *new CEO* denote a CEO who has never worked. We assume that there is never any scarcity of potential new CEOs and a share p_L of them is of the high type. The type of new CEOs is unobservable. Let a *successful CEO* denote a CEO which has already been hired at least once and completed a period of time T (which is now the "standard" contract duration). We denote by $p_H \equiv 1 - \kappa$ the probability that a successful CEO remains a high type. We assume that the type persistence is sufficiently large so that $p_H > p_L$. Finally, let a *failed CEO* denote a CEO who was hired and then fired.

[16] For simplicity, we assume that CEOs can only move to a different firm after their contract term T. Without loss of generality, their contract may also be renewed at the same firm.

We consider a competitive market for managerial talent, where firms offer CEOs a wage w based on her performance. The wage w is fixed for the duration of the contract (or until the CEO is fired). Since there is no scarcity of new CEO's, the wage of new CEO's is set equal to their reservation value, which we normalize to 0. For the same reason, no firm ever hires a failed CEO. Consider now the successful CEOs. In steady state, the fraction of (previously) successful CEOs among all CEO hires is given by

$$\mu = p_L(1-\mu) + p_H\mu = \frac{p_L}{1-p_H+p_L}.$$

In line with the intuition developed in Proposition 9, successful CEOs will receive a positive wage $\bar{w} > 0$ and they will be hired by the share μ of most productive firms. In particular, we obtain the following result, proven for the case where ρ is sufficiently large (firms are sufficiently myopic):

Proposition 10 *Assume ρ is sufficiently large. In steady state, there exists a cutoff $\bar{\pi}$ such that:*

- *Firms with productivity $\pi_t > \bar{\pi}$ hire only successful CEOs and pay a wage differential $\bar{w} > 0$.*
- *Firms with $\pi_t < \bar{\pi}$ hire only new CEOs.*
- *No firm hires failed CEOs.*
- *Firms at $\pi_t = \bar{\pi}$ are indifferent between hiring a new CEO or a (more expensive) successful CEO.*
- *Each firm's performance at even CEO transition times follows a Markov chain: if the firm is at level π_k, the probability of going up (down) one level is given by p_i^2 $((1-p_i)^2)$, where*

$$p_i \begin{cases} = p_L & \text{if } \pi_k < \bar{\pi} \\ \in [p_L, p_H] & \text{if } \pi_k = \bar{\pi} \\ = p_H & \text{if } \pi_k > \bar{\pi} \end{cases}$$

Proof. We prove by contradiction. Let $q = (q_l)_{l\in\mathbb{N}}$ be the steady state hiring profile where $q_l \in \{p_H, p_L\}$ denotes the type of manager hired at performance level $\pi_l \in \Pi$. Assume that our proposition does not hold. Then there must exists an m such that $q(m) = p_H$ but $q(m+1) = p_L$. But by Proposition 9, this is impossible. ■

When $\pi_k = \bar{\pi}$, the firm is indifferent over whether to hire a successful CEO or a new one. This creates (local) equilibrium multiplicity, which is allowed for in the statement of the proposition. For performance levels above and below $\bar{\pi}$, the stochastic process is uniquely defined.

6.3 Implications of the Endogenous Wage Model

Section 5 discussed the testable implications of the baseline model where CEOs can only work for one period. Let us now examine the additional predictions we can make when CEOs work for multiple periods and wages are endogenous.

In the equilibrium in Proposition 10, CEO careers display certain patterns. Bad CEOs are employed only once: after damaging the organizational capital of one firm, they become unemployable. Good CEOs are employed repeatedly and receive a compensation premium until they underperform. Firms with higher organizational capital hire better CEOs.

Proposition 11 *In steady state:*

(i) Firms with better performance and higher organizational capital employ CEOs of a better type, with better behavior, who are paid more.

(ii) The current employment status and compensation of a CEO depends on the change in performance and organizational capital of its previous firm.

(iii) A fixed-effect regression on data generated by this model returns significant individual coefficients but underestimates the true effect of individual CEOs on performance.

Proof. Parts (i) and (ii) are immediate consequences of Proposition 10.

For (iii), note that if a CEO is employed by n firms, she must perform well in the first $n-1$ firms and badly in the last one. Let us express performance changes in terms of levels, so the effect of a good CEO is 1 and the effect of a bad one is -1. The *true* fixed effect of a CEO with n is $\frac{n-2}{n}$.

Note, however, that a fixed-effect regression would attribute some of the CEO fixed effect to the firm. Consider a panel regression that includes the last N periods of every firm. Let $\bar{k}$ the performance level corresponding to $\bar{\pi}$. All firms whose initial performance level is $\bar{k} + N$

or higher will only hire CEOs with p_H. All firms whose initial performance level is $\bar{k} - N$ or lower will only hire CEOs with p_L. The average fixed effect difference between firms in the former set and firms in the latter set with

$$p_H - (1 - p_H) - (p_L - (1 - p_L)) = 2(p_H - p_L).$$

As the true fixed-effect of firms is zero, this means that the regression will underestimate the fixed effect of CEOs hired by high-performance firms and overestimate that of CEOs hired by low-performance firms. ■

Prediction (i) relates to an influential prediction of the CT literature: larger firms should hire better CEOs on average (Tervio (2008) and Gabaix and Landier (2008)).

Prediction (ii) relates the career path of CEOs to their effect on previous firms they worked for. Past employers of CEOs who currently command higher wages and work for more productive firms have experienced unusually strong growth in both performance and organizational capital.

Prediction (iii) relates to the estimation of CEO fixed effects developed by Bertrand and Schoar (2003). Consider an econometrician who observes the last N periods of a random sample of firms and estimates fixed effects for firms and CEOs. As firms with high organizational capital hire better CEOs on average, part of the CEO fixed effect will be attributed to the firm, thus underestimating the true causal effect of CEO on performance.

7 Conclusions

This paper began by noting that economists have studied the effect of management on firm performance from three distinct perspectives: CT, OC, and LC. The goal of the paper was to develop the most parsimonious model that can reconcile key patterns predicted or observed by the three perspectives. The main novel ingredient of the model was organizational capital, a set of productive assets that can only be produced with the direct input of the firm's leadership and it is subject to an agency problem. Besides yielding predictions that are consistent with the three perspectives, the model also generates novel predictions that combine OC and LC variables.

Appendix I: Proofs

Proof of Proposition 2: *In a steady state reachable from below, the mass of firms transitioning at performance level k is given by*

$$f^*(k) = M * \tfrac{(1-2p)^2}{(1-p)p} * k\left(\frac{p}{1-p}\right)^k$$

where the total mass of transitioning firms, given by $M = \frac{B}{\gamma^*} \equiv B\left(\frac{1-(1-2p)^2}{(1-2p)^2}\right)$

Proof: We proceed in two parts

Part 1: Linear Difference Equation: $\gamma \leq \gamma^*$ In steady state, $f(k)$ must satisfy the difference equation (6),

$$f(k) = (1+\gamma)\left(p^2 f(k-1) + (1-p)pf(k) + (1-p)^2 f(k+1)\right), \tag{7}$$

with the following boundary conditions:

$$f(0) = 0 \text{ and } f(1) = \frac{B}{(1+\gamma)(1-p)^2}$$

For every value of $\gamma = B/M$, standard techniques show that the difference equation (6) has at most one solution with non-negative values of $f(\cdot)$ as follows:

$$f(k) = \frac{A^k - D^k}{C},$$

where

$$\begin{aligned}
A &= A(\gamma) = \frac{1}{2}\left(\frac{1-2p(1+\gamma)+2p^2(1+\gamma)}{(1-p)^2(1+\gamma)} + \sqrt{\frac{1-4p(1+\gamma)+4p^2(1+\gamma)}{(1-p)^4(1+\gamma)^2}}\right); \\
D &= D(\gamma) = \frac{1}{2}\left(\frac{1-2p(1+\gamma)+2p^2(1+\gamma)}{(1-p)^2(1+\gamma)} - \sqrt{\frac{1-4p(1+\gamma)+4p^2(1+\gamma)}{(1-p)^4(1+\gamma)^2}}\right); \\
C &= C(\gamma) = \frac{1}{B}\frac{(1-p)^2}{1-(1-2p)^2}\sqrt{\frac{1-4p(1+\gamma)+4p^2(1+\gamma)}{(1-p)^4(1+\gamma)^2}}
\end{aligned}$$

Let

$$\gamma^* \equiv \frac{(1-2p)^2}{1-(1-2p)^2}$$

Consider the term under the three square roots that appears in the expressions of A, D, and C. When $\gamma > \gamma^*$, the term is negative, in which case it can be shown that $f(k)$ is strictly negative for certain values of k.[17] When $\gamma \to \gamma^*$, the expression above tends to:

$$\begin{aligned} f^*(k) &= B\left(\frac{1-(1-2p)^2}{(1-p)^2}\right)k\left(\frac{p}{1-p}\right)^{k-1} \\ &= M\tfrac{(1-2p)^2}{(1-p)^2}k\left(\frac{p}{1-p}\right)^{k-1} \end{aligned}$$

For every value $\gamma \leq \gamma^*$, the value of $f(k)$ is always positive and thus $f(k)$ is a potential steady state.

Lemma 2 *If a steady state exists, then it must be that $\gamma = B/M \leq \gamma^*$ and*

$$f(k) = \frac{A(\gamma)^k - G(\gamma)^k}{C(\gamma)},$$

Part 2: Transition Matrix: $\gamma \geq \gamma^*$. We want to characterize possible steady states of the recurrence relation

$$f_{r,s+1}(k) = \left(1 + \frac{B}{M_\tau}\right)\left[p^2 f_{r,s}(k-1) + 2p(1-p)f_{r,s}(k) + (1-p)^2 f_{r,s}(k+1)\right] \tag{8}$$

where M_τ is the total mass of firms transitioning at calendar time $\tau \equiv \tau(r,s,k)$. Given Assumption S1,

$$f_{r,s}(0) = 0 \text{ for all } s = 0, 1, ..., n$$

and we impose the following initial (or boundary) conditions:

$$\begin{aligned} f_{r,0}(k) &= 0 \text{ for all } k \neq 1 \\ f_{r,0}(k) &= M_0 \text{ for } k = 1 \end{aligned}$$

We now show that the only possible steady state is $f^*(\cdot)$.

Consider first an alternative recurrence relation, denoted by $f_{r,s}(k, N)$ where there are a finite number of performance levels $k \in \{1, 2, ..., N\}$ with N arbitrary large:

[17] The fact that the value under the square root is negative is not a problem per se because all terms with an even power drop out. However, for k large enough, $f(k) < 0$.

A feasible solution for $f(k)$ does not exist when γ is too high, because a very high birth rate leads to explosive growth in the number of firms.

$$f_{r,s+1}(k,N)=\left(1+\frac{B}{M_{N,\tau}}\right)\left(p^2 f_{r,s}(k-1,N)+2(1-p)\,p f_{r,s}(k,N)+(1-p)^2 f_{r,s}(k+1,N)\right)$$

for $k=1,2,...,N-1$, and

$$f_{r,s+1}(N,N)=\left(1+\frac{B}{M_{N,\tau}}\right)(p^2 f_{r,s}(N-1,N)+2(1-p)\,p f_{r,s}(N,N)),$$

where $M_{N,\tau}$ is the total mass of firms transitioning at calendar time $\tau\equiv\tau(r,s+1,k)$ (defined above). We impose the same boundary conditions as for our original recurrence relation: $f_{r,s}(0,N)=0$ for all s, $f_{r,0}(k,N)=0$ for all $k\neq 1$, and $f_0(k,N)=B$ for $k=1$.

The following result holds:

Lemma 3 *Assume that $f_{r,s}(k,N)$ converges to $f(k,N)>0$ but finite, then it must be that*

$$\gamma_N\equiv\frac{B}{\sum_k f(k,N)}\geq\gamma^*=\frac{(1-2p)^2}{4p(1-p)}$$

Proof. Assume that $f_{r,s}(k,N)$ converges to $f(k,N)>0$ but finite, then $\gamma_{N,\tau}\equiv B/M_{N,\tau}$ converges to

$$\gamma_N=B/\sum\nolimits_k f(k,N)$$

Define $\boldsymbol{f}_{\mathbf{N}}=[f(1,N),f(2,N),...,f(N,N)]^T$ and

$$\mathbf{A_N}=\begin{bmatrix} b & c & 0 & \cdots & 0\\ a & b & c & \cdots & 0\\ & \cdots & \cdots & \cdots & \\ 0 & \cdots & a & b & c\\ 0 & \cdots & 0 & a & b\end{bmatrix}_{N\times N}$$

with $a=(1+\gamma_N)\,p^2$, $b=2(1+\gamma_N)\,p(1-p)$, $c=(1+\gamma_N)(1-p)^2$. If there is convergence to $\boldsymbol{f}_{\mathbf{N}}$, we must have that

$$\boldsymbol{f}_{\mathbf{N}}=\mathbf{A_N}\times\boldsymbol{f}_{\mathbf{N}}$$

Let λ be the largest eigenvalue of $\mathbf{A_N}$. As the value of a cosine can never be larger than one, Theorem 16 in Cheng (2003)[18] implies that for a finite N, the largest eigenvalue of $\mathbf{A_N}$ is bounded above by

$$b+2\sqrt{ac}=4(1+\gamma_N)\,p(1-p)$$

[18] Cheng, Sui Sun. Partial difference equations. Vol. 3. CRC Press, 2003.

If $\lambda < 1$, there exists no vector $\boldsymbol{f} > \mathbf{0}$ such that $\boldsymbol{f} = \mathbf{A_N} \times \boldsymbol{f}$. Hence, a necessary condition for $f_{r,s}(k,N)$ to converge to a positive steady state (that is $f(k,N) > 0$ but finite) is that $\lambda \geq 1$ or still $4(1+\gamma_N)\,p(1-p) \geq 1$ or still

$$\gamma_N \geq \frac{(1-2p)^2}{4p(1-p)}$$

■

Consider now again our original recurrence relation $f_{r,s}(k)$ and denote $\gamma_\tau \equiv B/M_\tau$. Given the initial conditions for $f_{r,0}(k)$, we have that

$$f_{r,s}(k) = 0 \text{ for } k > s \tag{9}$$

It follows that for any $s > 0$

$$\lim_{N\to\infty} f_{r,s}(k,N) = f_{r,s}(k) \tag{10}$$

and thus also $\lim_{N\to\infty} \gamma_{N,\tau} = \gamma_\tau$.

Assume now that $f_{r,s}(k)$ converges to $f(k) > 0$ but finite. Then from (9) and (10), for N large enough, also $f_{r,s}(k,N)$ converges to $f(k,N) > 0$ and $\gamma_{N,\tau}$ converges to $\gamma_N > 0$ and

$$\lim_{N\to\infty} f(k,N) = f(k) \text{ and } \lim_{N\to\infty} \gamma_N = \gamma$$

Together with the previous lemma this implies:

Lemma 4 *Assume that $f_{r,s}(k)$ converges to $f(k) > 0$ but finite, then it must be that*

$$\gamma = B/\sum\nolimits_k f(k) \geq \gamma^* = \frac{(1-2p)^2}{4p(1-p)}$$

We conclude that if a steady state $f(k)$ exists, then it must that

$$\frac{B}{M} = \gamma^* \equiv \frac{(1-2p)^2}{4p(1-p)}$$

The linear difference equation (6) then implies that

$$f(k) = f^*(k) \equiv M * \frac{(1-2p)^2}{(1-p)^2} k \left(\frac{p}{1-p}\right)^{k-1}$$

This concludes the proof of Proposition 2.

Proof of Corrolary 3. *In steady state, for $\Omega_k \in \Pi$ large, $\tilde{\phi}(\Omega_k)$ approximates a power law: There exists a $c > 0$ such that for Ω_k large $\tilde{\phi}(\Omega_k) \approx c * \Omega_k^{-\zeta}$ with $\zeta = h * \ln \frac{1-p}{p}$ with $h \equiv 1/(2\ln(1+\Delta))$.*

Proof. From Proposition 2, we have that $f(\Omega_k) \sim k\left(\frac{p}{1-p}\right)^k$ for $\Omega_k \in \Pi$. Wlog, set $\Omega_0 = 1$. Since $\Omega_k = (1+\Delta)^{2k}$, we can rewrite this as

$$f(\Omega_k) \sim h * \ln \Omega_k * \left(\frac{p}{1-p}\right)^{h \ln \Omega_k}$$

where $h = 1/(2\ln(1+\Delta)) > 0$. Consider now $\Omega_{k+l} = a\Omega_k \in \Pi$ where $a = (1+\Delta)^{2l}$. Then

$$f(a\Omega_k)/f(\Omega_k) = \ln a\Omega_k / \ln \Omega_k * \left(\frac{p}{1-p}\right)^{h(\ln a\Omega_k - \ln \Omega_k)}$$

from which

$$\lim_{k\to\infty} f(a\Omega_k)/f(\Omega_k) = a^{-\zeta}$$

with $\zeta = -\frac{1}{2\ln(1+\Delta)} * \ln \frac{p}{1-p}$. It follows that for Ω_k large, $f(\Omega_k) \approx c * \Omega_k^{-\zeta}$ for some constant c.
■

Proof of Proposition 4. *Let $\Phi(\Omega')$ be the cumulative distribution of all firms (transitioning and non-transitioning) with organizational capital Ω smaller or equal than Ω'. Then $\tilde{\Phi}(k-1) < \Phi(\Omega_k) < \tilde{\Phi}(k)$ with $\Phi(\Omega_k) = \tilde{\Phi}(k-1) + \tilde{\phi}(k)(1-p)\frac{\bar{t}+(1-p)\bar{t}+pT}{2((1-p)\bar{t}+pT)}$*

Proof. Consider the cumulative mass of all firms – transitioning and non-transitioning – whose organizational capital $\Omega \in \mathbb{R}^+$ is smaller or equal than $\Omega_k = \pi_k \in \Pi^+$. Denoting this cumulative mass by $G(\Omega_k)$, we have that

$$G(\Omega_1) = f(1)(1-p)\left[\bar{t} + (1-p)\bar{t} + pT\right]$$

and

$$\begin{aligned} G(\Omega_k) &= G(\Omega_{k-1}) \\ &\quad + f(k-1)p\left[T + pT + (1-p)\bar{t}\right] \\ &\quad + f(k)(1-p)\left[\bar{t} + (1-p)\bar{t} + pT\right] \end{aligned}$$

Some tedious algebra then yields that

$$\begin{aligned} G(\Omega_k) &= 2\left(pT + (1-p)\bar{t}\right)\sum\nolimits_{m<k} f(m) + f(k)(1-p)\left[\bar{t} + (1-p)\bar{t} + pT\right] \\ &= \bar{M} * F(k-1) + f(k)(1-p)\left[\bar{t} + (1-p)\bar{t} + pT\right] \end{aligned}$$

It follows that $\Phi(\Omega_k) \equiv G(\Omega_k)/\bar{M}$, the cumulative distribution of all firms (transitioning and non-transitioning) with organizational capital $\Omega \leq \Omega_k$, is given by

$$\begin{aligned}\Phi(\Omega_k) &= \tilde{\Phi}(k-1) + f(k)(1-p)\frac{[\bar{t} + (1-p)\bar{t} + pT]}{\bar{M}} \\ &= \tilde{\Phi}(k-1) + \phi(k)(1-p)\frac{\bar{t} + ((1-p)\bar{t} + pT)}{2((1-p)\bar{t} + pT)}\end{aligned}$$

Alternatively, we can write

$$G(\Omega_k) = \bar{M} * \tilde{\Phi}(k) - f(k)p\,[T + pT + (1-p)\bar{t}]$$

and thus

$$\Phi(\Omega_k) = \tilde{\Phi}(k) - \phi(k)p\frac{T + (pT + (1-p)\bar{t})}{2((1-p)\bar{t} + pT)}$$

We conclude that $\tilde{\Phi}(k-1) < \Phi(\Omega_k) < \tilde{\Phi}(k)$.

■

Proof of Proposition 9. *Consider a CEO with p' and a CEO with $p'' > p'$. Let Ω' and Ω'' denote the organizational capital levels of the firms employing the two CEOs respectively. If firms are sufficiently impatient (ρ is sufficiently high), in steady state $\Omega' \leq \Omega''$.*

Proof. Suppose for contradiction that $\Omega' > \Omega''$.

Let $W(p_j)$ represent the expected discounted cost given the instantaneous wage w_j and the probability of success p_j of employing a CEO of category j. Note that $W(p_j)$ is independent of the organizational capital of the firm that employs the CEO.

Let u_k denote the steady state expected discounted payoff of a firm at level k (who does not yet know the quality of its new CEO). The payoff of a firm at level k who hires a CEO of category p is given by

$$\begin{aligned}\tilde{u}_k(p) &= p\left(\int_0^T e^{-\rho t}e^{(\theta^H-\delta)t}\Omega_k dt + e^{-\rho T}u_{k+1}\right) + (1-p)\left(\int_0^{\bar{t}} e^{-\rho t}e^{(\theta^H-\delta)t}\Omega_k dt + e^{-\rho\bar{t}}u_{k-1}\right) \\ &= p\left(\frac{1 - e^{-\rho T}e^{(\theta^H-\delta)T}}{\rho+\delta-\theta^H}\Omega_k + e^{-\rho T}u_{k+1}\right) + (1-p)\left(\frac{1 - e^{-\rho\bar{t}}e^{(\theta^H-\delta)\bar{t}}}{\rho+\delta-\theta^H}\Omega_k + e^{-\rho\bar{t}}u_{k-1}\right) \\ &= v_k + e^{-\rho\bar{t}}z_k(p)\end{aligned}$$

where

$$
\begin{aligned}
v_k(p) &= \frac{1-e^{-\rho\bar{t}}e^{\left(\theta^H-\delta\right)\bar{t}}}{\rho+\delta-\theta^H}\Omega_k \\
z_k\left(p\right) &= p\left(\frac{1-e^{-\rho(T-\bar{t})}e^{\left(\theta^H-\delta\right)(T-\bar{t})}}{\rho+\delta-\theta^H}\Omega_k e^{\left(\theta^H-\delta\right)\bar{t}}\right)+(1-p)\left(\frac{1-e^{-\rho(T-\bar{t})}e^{\left(\theta^H-\delta\right)(T-\bar{t})}}{\rho+\delta-\theta^H}\Omega_k e^{-\delta\bar{t}}\right) \\
&\quad +e^{-\rho(T-\bar{t})}(pu_{k+1}+(1-p)u'_{k-1})
\end{aligned}
$$

where u'_{k-1} is defined as the expected steady state discounted payoff of a firm who $T-\bar{t}$ periods had organizational capital Ω_{k-1}. Note that we necessarliy must have that $u'_{k-1} < u_k$.

It is optimal for a firm at level k to employ a CEO with p' rather than one with p'' if

$$\tilde{u}_k\left(p'\right)-W\left(p'\right)\geq\tilde{u}_k\left(p''\right)-W\left(p''\right)$$

Conversely, it is optimal for a firm at level m to employ a CEO with p'' rather than one with p' if

$$\tilde{u}_m\left(p'\right)-W\left(p'\right)\leq\tilde{u}_m\left(p''\right)-W\left(p''\right)$$

Subtracting one condition from the other we obtain

$$\tilde{u}_m\left(p'\right)-\tilde{u}_k\left(p'\right)\leq\tilde{u}_m\left(p''\right)-\tilde{u}_k\left(p''\right)$$

which can be re-written as

$$z_m\left(p'\right)-z_k\left(p'\right)\leq z_m\left(p''\right)-z_k\left(p''\right) \tag{11}$$

Note that

$$
\begin{aligned}
\lim_{\rho\to\infty}\left(\rho+\delta-\theta^H\right)z_k\left(p\right) &= p\Omega_k e^{\left(\theta^H-\delta\right)\bar{t}}+(1-p)\Omega_k e^{-\delta\bar{t}} \\
&= \Omega_k e^{-\delta\bar{t}}+p\Omega_k(e^{\left(\theta^H-\delta\right)\bar{t}}-e^{-\delta\bar{t}})
\end{aligned}
$$

Thus,

$$\lim_{\rho\to\infty}z_k\left(p\right)=\frac{\Omega_k}{\rho+\delta-\theta^H}e^{-\delta\bar{t}}+p\frac{\Omega_k}{\rho+\delta-\theta^H}\left(e^{\left(\theta^H-\delta\right)\bar{t}}-e^{-\delta\bar{t}}\right)$$

For ρ large enough, inequality (11) holds if and only if

$$p^{'}\Omega_m-p^{'}\Omega_k\leq p''\Omega_m-p''\Omega_k$$

namely $\left(p^{'}-p''\right)(\Omega_m-\Omega_k)\leq 0$, which is false when $p^{'}<p''$ and $\Omega_m>\Omega_k$. ■

Appendix II: Full Agency Problem

We keep the model defined in Section 2 except for the following modifications:

- The agent receives a minimum wage $w > 0$ while employed. The wage is instantaneous and it is a share of the company's performance when the agent is hired (this assumption is made to abstract from a scale effect). The wage can be thought of as $w = \bar{w} + \psi$, where $\bar{w}$ is a minimum statutory monetary wage and ψ is a psychological benefit of being a CEO. As the firm owner must pay $\bar{w}$ to all CEOs and the firm must always have a CEO, the minimum wage can be omitted when solving the firm-owners dynamic optimization problem.

- The firm owner can also promise a performance bonus to the CEO. The bonus may depend on performance as well as any message that the agent may send.

- The CEO and the firm owner have the same discount rate ρ.

We say that a contract is first-best contract if it guarantees that the firm is always run by a good CEO.

Proposition 12 *There exists a contract that achieves first best. However, for any positive w, if p is sufficiently small, the firm will not offer it.*

Proof. In order to achieve an efficient outcome, the owner must induce bad CEOs to resign as soon as they are hired – or equivalently, reveal their type truthfully and be fired. Suppose the owner offers a performance bonus b if a CEO resigns right after being hired. If a bad CEO does not resign at zero, he receives payoff

$$\int_t^{t+\bar{t}} e^{-\rho t} w dt = \frac{1 - e^{-\rho \bar{t}}}{\rho} w \Omega_t.$$

If he resigns (and we assume that any other bad CEO resigns immediately), he instead gets b. Thus, the minimum cost for the principal (evaluated at the beginning of the relationship) for persuading one bad CEO to resign (which satisfies the incentive constraint) is

$$b = \frac{1 - e^{-\rho \bar{t}}}{\rho} w \Omega_t$$

Note that given a bonus b at time 0, a good CEO strictly prefers not to resign as her tenure at the firm, T, is longer than that of a bad CEO, $\bar{t}$.

If the owner gets rid of a bad CEO, she still faces a probability $1-p$ that the next CEO is bad as well, implying that she would have to pay b again. Thus, the average cost of guaranteeing that the CEO hired at t is good for sure is:

$$\left(1-p+(1-p)^2+...\right)\frac{1-e^{-\rho\bar{t}}}{\rho}w\Omega_t=\left(\frac{1-p}{p}\right)\frac{1-e^{-\rho\bar{t}}}{\rho}w\Omega_t.$$

We now compare the expected value of a firm at t who chooses to implement the incentive scheme above as compared to one that does not (and therefore behaves like the firm in Proposition 1). With the incentive scheme, all CEOs are good and have a tenure of length T . At each CEO transition, the firm sustains expected cost $\frac{1-p}{p}\frac{1-e^{-\rho\bar{t}}}{\rho}w\Omega_t$. The expected value of the firm is given by

$$\begin{aligned}
\tilde{V}_t &= \left(\frac{1}{\rho+\delta-\theta^H}\left(1-e^{-\left(\rho+\delta-\theta^H\right)T}\right)-\left(\frac{1-p}{p}\right)\frac{1-e^{-\rho\bar{t}}}{\rho}w\right)\Omega_t+\tilde{V}_{t+T}\\
&= \Omega_t\left(\frac{1}{\rho+\delta-\theta^H}\left(1-e^{-\left(\rho+\delta-\theta^H\right)T}\right)-\left(\frac{1-p}{p}\right)\frac{1-e^{-\rho\bar{t}}}{\rho}w\right)\sum_{k=0}^{\infty}e^{-\left(\rho+\delta-\theta^H\right)Tk}\\
&= \Omega_t\left(\frac{1}{\rho+\delta-\theta^H}\left(1-e^{-\left(\rho+\delta-\theta^H\right)T}\right)-\left(\frac{1-p}{p}\right)\frac{1-e^{-\rho\bar{t}}}{\rho}w\right)\frac{1}{1-e^{-\left(\rho+\delta-\theta^H\right)T}}
\end{aligned}$$

Instead, as we know from Proposition 1, the value of a firm that does not offer this incentive scheme is

$$\Omega_t\frac{1}{\rho+\delta-\theta^H}\frac{1-pe^{-\left(\rho+\delta-\theta^H\right)T}-(1-p)\,e^{-\left(\rho+\delta-\theta^H\right)\bar{t}}}{1-pe^{-\left(\rho+\delta-\theta^H\right)T}-(1-p)\,e^{-(\rho+\delta)\bar{t}}}$$

The owner does not find it in her interest to induce bad CEOs to resign if

$$\begin{aligned}
&\left(\frac{1-p}{p}\right)\frac{1-e^{-\rho\bar{t}}}{\rho}w\\
\geq\;&\left(\frac{1}{\rho+\delta-\theta^H}-\frac{1}{\rho+\delta-\theta^H}\frac{1-pe^{-\left(\rho+\delta-\theta^H\right)T}-(1-p)\,e^{-\left(\rho+\delta-\theta^H\right)\bar{t}}}{1-pe^{-\left(\rho+\delta-\theta^H\right)T}-(1-p)\,e^{-(\rho+\delta)\bar{t}}}\right)\left(1-e^{-\left(\rho+\delta-\theta^H\right)T}\right)\\
=\;&\frac{1}{\rho+\delta-\theta^H}\frac{(1-p)\left[e^{-\left(\rho+\delta-\theta^H\right)\bar{t}}-e^{-(\rho+\delta)\bar{t}}\right]}{1-pe^{-\left(\rho+\delta-\theta^H\right)T}-(1-p)\,e^{-(\rho+\delta)\bar{t}}}\left(1-e^{-\left(\rho+\delta-\theta^H\right)T}\right)
\end{aligned}$$

That is

$$w\geq p\frac{\rho}{\rho+\delta-\theta^H}\frac{1-e^{-\left(\rho+\delta-\theta^H\right)T}}{1-e^{-\rho\bar{t}}}\frac{e^{-\left(\rho+\delta-\theta^H\right)\bar{t}}-e^{-(\rho+\delta)\bar{t}}}{1-pe^{-\left(\rho+\delta-\theta^H\right)T}-(1-p)\,e^{-(\rho+\delta)\bar{t}}}$$

from which we can see the statement of the Proposition. ■

The intuition for this result is that, in order to achieve an efficient outcome, the owner must induce bad CEOs to resign as soon as they are hired – or equivalently, reveal their type truthfully and be fired. As such, the firm must offer the bad CEO an incentive scheme that pays at least as much as what a bad CEO would get by staying at the firm for $\bar{t}$. This compensation must be paid to all the bad CEOs that are hired and resign immediately. The latter part grows unboundedly as $p \to 0$.

References

Bandiera, Oriana, Stephen Hansen, Andrea Prat, and Raffaella Sadun (2016). "CEO Behavior and Firm Performance." working paper.

Bebchuk, Lucian (2009). Pay Without Performance: The Unfulfilled Promise of Executive Compensation. Harvard University Press.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter (2016). Management practices, workforce selection and productivity. No. w22101. National Bureau of Economic Research.

Bennedsen, Morten, Kasper Meisner Nielsen, Francisco Pérez-González, and Daniel Wolfenzon (2007). "Inside the family firm: The role of families in succession decisions and performance." The Quarterly Journal of Economics 122, no. 2: 647-691.

Bertrand, Marianne (2009). "CEOs." Annu. Rev. Econ. 1.1 : 121-150.

Bertrand, Marianne and Antoinette Schoar (2003). Managing with style: The effect of managers on firm policies. The Quarterly Journal of Economics, 1169-1208.

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., and Van Reenen, J. (2016). What Drives Differences in Management?. Stanford mimeo.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013). "Does Management Matter? Evidence from India." Quarterly Journal of Economics 128, no. 1

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2016). "Management as a Technology?". No. w22327. National Bureau of Economic Research, 2016

Bloom, Nicholas, and John Van Reenen (2007). "Measuring and Explaining Management Practices Across Firms and Countries." Quarterly journal of economics 122 (4): 1351-1408.

Board, Simon, Moritz Meyer-ter-Vehn, and Tomasz Sadzik. "Recruiting Talent." (2016). Working paper.

Chassang, Sylvain (2010) "Building routines: Learning, cooperation, and the dynamics of incomplete relational contracts." The American Economic Review 100.1: 448-465.

Drucker, Peter (1967), The Effective Executive, New York: Harper & Row

Ellison, Glen and Richard Holden (2014). "A Theory of Rule Development." Journal of Law, Economics and Organization, 30(4), 649-682.

Ericson, Richard, and Ariel Pakes (1995). "Markov-perfect industry dynamics: A framework for empirical work." The Review of Economic Studies 62.1: 53-82.

Gabaix, Xavier, and Augustin Landier (2008). "Why has CEO pay increased so much?." The Quarterly Journal of Economics 123, no. : 49-100.

Garicano, Luis, and Esteban Rossi-Hansberg (2006). "Organization and Inequality in a Knowledge Economy." The Quarterly Journal of Economics 121, no. 4, 1383–1435.

Gibbons, Robert, and Rebecca Henderson (2013). "What Do Managers Do? Exploring Persistent Performance Differences among Seemingly Similar Enterprises." Handbook of Organizational Economics, eds. Robert Gibbons and John Roberts, Princeton University Press, 680-731.

Graham, John R., Campbell R. Harvey, Jillian A. Popadak, and Shivaram Rajgopal (2016). "Corporate culture: Evidence from the field." working paper.

Halac, Marina, and Andrea Prat (2016). "Managerial attention and worker performance." The American Economic Review 106, no. 10: 3104-3132.

Hannes et a. (2010).

Hermalin, Benjamin E (2013). "Leadership and Corporate Culture." The Handbook of Organizational Economics, eds. Robert Gibbons and John Roberts, Princeton University Press: 432-478.

Hopenhayn, Hugo A. (1992) "Entry, exit, and firm dynamics in long run equilibrium." Econometrica, 60(5): 1127-1150.

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi (1997). "The effects of human resource management practices on productivity: A study of steel finishing lines." The American Economic Review: 291-313.

Kaplan, Steven N., Mark M. Klebanov, and Morten Sorensen (2012). "Which CEO characteristics and abilities matter?." The Journal of Finance 67.3: 973-1007.

Kotter, John P. (2001). "What Leaders Really Do." Harvard Business Review, Vol. 79(5): 62-70.

Li, Jin, Niko Matouschek and Michael Powell (2017). "Power Dynamics in Organizations." American Economic Journal: Microeconomics, 9(1): 217-41.

Lucas Jr, Robert E. (1978). On the size distribution of business firms. The Bell Journal of Economics, 508-523

Milgrom, Paul, and John Roberts (1995). "Complementarities and fit strategy, structure, and organizational change in manufacturing." Journal of accounting and economics 19, no. 2 (1995): 179-208.

Powell, Michael (2016). "Productivity and credibility in industry equilibrium. Working paper.

Schein, Edgar (2010). Organizational Culture and Leadership. John Wiley & Sons.

Stein, Jeremy C.(1989) "Efficient capital markets, inefficient firms: A model of myopic corporate behavior." The Quarterly Journal of Economics 104.4: 655-669.

Syverson, Chad (2004). "Product Substitutability and Productivity Dispersion." Review of Economics and Statistics, 86(2): 534-50.

Syverson, Chad (2011) "What determines productivity?." Journal of Economic literature 49.2: 326-365.

Terviö, Marko (2008). The difference that CEOs make: An assignment model approach. The American Economic Review, 98(3), 642-668.

Von Thadden, Ernst-Ludwig (1995). "Long-term contracts, short-term investment and monitoring." The Review of Economic Studies 62.4: 557-575.