Board of Directors' Theories and Choice of CEO

Arnaldo Camuffo[*], Miguel Espinosa[†],
Alfonso Gambardella[‡] & Andrea Pignataro[§]

October 24, 2023

Abstract

This paper develops a model showing that Board of Directors that envision theories of long-term firm growth spend more time identifying the right CEO profile for their visions. These "CEO leaders" spend more time in complementary theories and long-term visions, triggering a powerful source of firm performance compared to firms that focus on the short-run, do not have good theories, and hire "CEO managers" with the same short-term perspective and focused on managing the status quo. In equilibrium, CEO managers negotiate a higher share of the firm's current output as compensation because, unlike CEO leaders, they cannot rely on higher future output. We provide examples consistent with the functioning, mechanisms, and implications of our model.

JEL: L21, L26, M13, M21

Keywords: Choice of CEO, Theory of the Firm, CEO Compensation

[*]Bocconi University, ION Management Science Lab, arnaldo.camuffo@unibocconi.it.
[†]Bocconi University, CEPR, CESIfo, miguel.espinosa@unibocconi.it.
[‡]Bocconi University, ION Management Science Lab, CEPR, alfonso.gambardella@unibocconi.it.
[§]ION Group, ION Management Science Lab, ap@iongroup.com

1 Introduction

New events or the drive to embrace opportunities often push companies to envision their future and seek novel and valuable paths forward. These decisions are highly impactful, but also uncertain and idiosyncratic, as Board of Directors (BoDs) and executives cannot rely on past data or useful analogies to implement them (Choi and Levinthal, 2022).

We follow recent research at the crossroads of strategy and decision science that posits that BoDs and executives can make these *low-frequency high-impact* (LFHI) strategic decisions (Camuffo et al., 2023a) by developing theories that allow them to envision future states and better understand their causes and implications. A theory is an abstract, causal representation of the world that, through conjectures, leads to the formulation of a future state space, along with beliefs of how likely are the future states that they envision, as well as how likely the theories they formulate are true (Nickerson and Zenger, 2004; Baer et al., 2013; Felin and Zenger, 2017; Ortoleva, 2012; Karni, 2022; Camuffo et al., 2023b).

Most often, these decisions are associated with changes in the company leadership, in that the BoD needs a CEO (more generally a top management team) to further develop and implement the change. This paper focuses on the case in which the BoD chooses the type of CEO that will implement their vision of the future of the BoD. We develop a model of CEO succession when BoDs form theories about their firms' future and the type of CEO that best matches them.

Our model shows that BoDs with greater ability to develop theories are more likely to choose higher expected growth paths and to profile CEOs that match their vision. At the same time, CEOs who match high-growth BoDs' theories spend relatively less time on execution to raise current performance and more time on contributing to developing, complementing, and implementing the BoD's theory. Moreover, in equilibrium, BoDs with theories and higher expected growth compensate CEOs with a lower share of a

larger expected future output. In brief, our model shows that theory development and implementation, high expected growth, CEO profiling, and a long-term horizon, are all complementary and they allow to attract CEOs who rely on their contribution to the development of a firm's theory to generate future returns.

Section 2 grounds the paper in the CEO succession literature and its relationship to corporate strategy and corporate governance. Section 3 provides a motivating example. Sections 4-6 develop the model. They articulate the elaboration and functioning of theories, and their implications for the choice and sorting of CEO across firms with different theories. Section 7 provides new examples that illustrate the model. Section 8 concludes. Appendix A and B provide details and extensions of the model.

2 Boards of Directors and Choice of CEO

The literature on CEO succession has been systematized by Finkelstein et al. (2009) and extensively reviewed in subsequent years (Cragun et al., 2016; Berns and Klarner, 2017; Schepker et al., 2017; Nyberg et al., 2021). It identifies four critical dimensions of CEO succession: (1) if and why it occurs, (2) how it occurs, (3) who is recruited, selected, and hired, and (4) what consequences it generates. This framework has been widely used to understand the antecedents, process, contingencies, and performance of CEO succession.

In this paper, we focus on the choice of CEO as a LFHI strategic decision interdependent with the other LFHI strategic decisions of the firm, especially those related to the envisioning of new markets and technologies (Elfenbein and Sterling, 2018). The paper relates to several streams of the literature on BoD and the choice of CEO.

First, it echoes the research that examines the extent to which CEOs affect firm performance (Finkelstein and Boyd, 1998; Hambrick and Quigley, 2014; Nguyen and Nielsen, 2014; Quigley and Graffin, 2017; Fitza, 2017; Quigley et al., 2017). These studies estimate

that CEOs can account for 10%-30% of a firm's financial performance, but also show that the degree of such impact can vary significantly. Moreover, they demonstrate that the proportion of variance in performance attributed to CEOs has steadily increased over the past few decades (Quigley and Graffin, 2017) and that shareholders perceive CEOs as increasingly crucial drivers of firm outcomes.

Second, it relates to classic studies on BoD succession planning (Zajac, 1990; Shen and Cannella Jr, 2003; Charan, 2005; Harrell, 2016; Cheng et al., 2020) and how succession plans can help BoDs cope with uncertainty regarding CEO succession, ultimately improving firm performance. For example, Schepker et al. (2018) found that formalized succession planning processes lead to more and higher-quality candidates.

Third, it draws upon the literature that investigates which CEO's characteristics and abilities matter for firm performance. For instance, Kaplan et al. (2012) found that CEOs' general and execution skills are positively linked to firm performance, while Waldman et al. (2001) and Agle et al. (2006) showed that some leadership attributes, such as transactional and charismatic leadership, are predictive of future financial performance, especially in uncertain environments. Falato et al. (2015) identified various reputational, career, and educational credentials that signal CEOs' skills, with higher credentials associated with better firm performance. Ou et al. (2018) revealed that CEO humility fosters collaboration, shared decision-making, and a shared vision, ultimately leading to increased firm performance. At the same time, overconfident CEOs make asymmetric decisions that adversely affect firm performance (Malmendier and Tate, 2008; Galasso and Simcoe, 2011; Kaplan et al., 2022).

Fourth, our paper offers insight into the reasons why CEOs might be over or underpaid in the market (as observed by Quigley et al., 2020) or at least why they might over or under-capture the value that they create through compensation (as highlighted by Nguyen

and Nielsen (2014)). Our model provides an explanation for why low-growth firms may overpay their CEOs in the short run, while high-growth firms do not.

Finally, our model aligns with the results reported by Bandiera et al. (2020), who observed that CEOs have diverse attitudes toward cultivating a vision and stimulating growth, with some of them adopting a leadership stance while others focus more on managing tasks. The former ("leader CEOs") usually dedicate more time to innovation and the creation of a company's vision, which can ultimately translate into higher productivity and long-term growth for the company. Their findings show that the BoD decision to hire a "leader CEO" rather than a "manager CEO" can have substantial effects on firm performance.

3 Toyota's Shift to Battery Electric Vehicles and new CEO appointment

To illustrate our model, we use the public announcement of Toyota in January 2023 of their desire to shift technologies from hybrid to fully electric vehicles and their corresponding CEO succession plan.

Since Akio Toyoda – grandson of the founder, Sijchiro Toyoda – became President of the company and Takeshi Uchiyamada its Chairman in 2009, the theory about the future of the auto industry of the BoD of Toyota was that multiple powertrain technologies, including hybrid vehicles (HEVs and PHEVs), hydrogen, fuel-cell, and battery electric vehicles (BEVs), would ensure the best fuel efficiency and potential carbon neutrality. In the absence of a clearly dominant technology, Toyota believed that the auto industry would evolve into an industry with segmented markets in which multiple types of powertrain technologies would cohabit.

This theory was grounded in Toyota's intuition – back in the early 1990s – that environmentally sustainable hybrids would replace internal combustion engines. Because of the scarcity of lithium and the suboptimal lithium battery technology, BEVs (fully electric vehicles) were not the best solution to reach fuel efficiency and carbon neutrality. Gill Pratt, Toyota's chief scientist, and former professor of Electrical Engineering and Computer Science at MIT, articulated the cornerstone assumption of this theory: "*putting a lot of lithium into large batteries for EVs is a waste of precious resources if drivers are going to use them mostly for fairly short commutes. The same amount of lithium could be used more efficiently in terms of cutting carbon emissions by dividing it among more hybrids (including plug-ins) with smaller batteries.*"[1] For the past 15 years, Toyota has grounded all its strategic decisions (R&D, new product development, manufacturing, supply chain, and distribution) on this theory. The accomplishments of the past 15 years strengthened Akio Toyoda's belief that his theory was solid.

This theory had a corresponding CEO/top management team theory, with a profile characterized by knowledge of hybrid technologies (to align with the corporate theory), seniority in the company (to ensure continuity), and competence in the established regional markets (to support profitability). Akio Toyoda had seen himself as the best match for such a profile and was appointed President and CEO of Toyota in 2009, with Takeshi Uchiyamada acting as Chairman. The rest of the top management team also had competencies –experience– rooted in hybrid technologies.

In 2022, Akio Toyoda questioned how the company and the industry would look in the future. He started to believe that BEVs might become the dominant powertrain technology. The key factors that prompted this revision of the theory were improvements in battery technology (the major constraint together with charging infrastructure, see

[1] https://newsroom.toyota.eu/toyotas-dr-gill-pratt-explains-the-advantage-of-toyotas-multi-technology-path-to-carbon-neutrality-at-davos-world-economic-forum/.

Murmann and Schuler, 2023), and pressure from regulators around the world. Both factors were likely to increase consumer demand and buoy the industry shift towards electrification. He also started to observe that other automotive manufacturers were laying out plans to go all-electric. General Motors planned to reach this goal by 2035, and Volkswagen was investing more than 100 billion dollars in EVs. Moreover, Tesla and the Chinese company BYD were growing rapidly as BEV-only car makers.

The new BEV theory of Toyota was more uncertain. There is still little data on BEVs and on all the contingencies necessary to make it the cornerstone of the future auto industry. The multiple-technology-based industry relied, instead, on more than two decades of HEV and PHEV successful design, production, and sales. For these reasons, the new theory also had a longer-term orientation. In Akio Toyoda's words:

> "*I believe that in terms of crisis, two paths appear before us. One is the path toward a short-term success or a quick victory [Theory 1, based on HEVs and PHEVs, which already proved to be successful]. The other is (...) to transform Toyota into a mobility company (...) to make ever-better cars and to become a best-in-town company, beloved and relied upon by stakeholders in each region around the world [Theory 2, exploratory, growth-oriented, more open to BEVs]*"[2]

The change in theory about the future of the firm occurred concomitantly with a change in theory about the future leadership. Akio Toyoda decided to retire from President of the Toyota Motor Corporation, starting the succession planning process. He articulated a theory of Toyota's CEO consistent with Toyota's newly envisioned future. The ideal candidate (the CEO profile) should embrace Toyota's values, be passionate about cars and about driving them, be knowledgeable about BEVs, have stamina and en-

[2]https://toyotatimes.jp/en/newscast/002.html

ergy ("be young") and a strong sense of pursuit (a "Master Driver" in Toyota's jargon) to take up the challenge and the long-term perspective implied by the company's paradigm shift.[3] In the previous profile the emphasis was on competence about the hybrid technology, now it is on competence about the BEV technology, but also on the CEO energy and a sense of pursuit to manage an important, difficult, and uncertain transition.

Toyota's BoD chose Koji Sato. Sato, 53 (unusually young by Toyota's standards), had joined Toyota Motor Corporation in 1992, and engaged in the development of parts and components of Corolla and Lexus since 1994. In 2016 he became Chief Engineer for Lexus and in 2020 he became President of Lexus International Company, where he led the design of the Lexus full electric RZ models and the GAZOO Racing Company.

4 Model Set-Up

4.1 Firm Output

Our model starts with an initial period in which a BoD develops a theory about the future of the firm –vision– and of the ideal CEO profile to implement it. We mark the end of this initial period as time 0, which is the time in which the CEO is hired and starts the tenure. We study the BoD development and choice of theories, the implications of these theories for the choice of CEO, the CEO behavior, and the implications of both BoD theories and CEO behavior for firm performance.

In making their choices, BoDs and CEOs look at the expected discounted sum of their future returns. For simplicity, we assume that both BoD and CEO expect that the tenure is long enough so is equivalent to an infinite discounted sum of future returns. The model can be extended to the case in which the CEO leaves the company, at the cost of

[3] https://toyotatimes.jp/en/newscast/002.html

additional complications that do not improve the insights gained from the balance case.

Let t be the time that, in the steady state, the CEO spends on executing the production of output in a given period. Let y be the output produced by the firm in the first period, that is at the end of time 1. The production function is $y = At^{\frac{1}{2}}$, where A is a productivity parameter and for simplicity and without loss of generality we abstract from the choice of other inputs.

There are two elements of randomness in this process. First, the BoD's theory predicts conditions for innovation and growth. However, if the theory is wrong, these conditions do not generate any growth. Second, if the theory is right, there is a probability λ that in any given period the firm realizes the potential growth rate $g > 0$ of A between any two consecutive periods. The probability λ is in the control of the CEO. We assume that $\lambda = z$, where z is the time spent by the CEO in developing complementary theories or activities that raise the probability of realizing the opportunities laid out by the BoD's theory.[4]

We normalize the total time available to the CEO in each period so that $t + z = 1$ and t and z represent the fraction of time spent in execution versus developing the vision, respectively. The production function is then $y = A(1 - z)^{\frac{1}{2}}$. We assume for simplicity that g becomes known during the first period. In the next section, this implies that the CEO chooses z knowing g. However, g is not known at time 0 when the BoD develops its theories, hires the CEO, and, as we will see in Section 6.2, BoD and CEO negotiate the CEO compensation.

[4]We assume that potential growth is either $g > 0$ or 0. The implications of the model do not change if we assume that g has a negative lower bound $g_0 \in (-1, 0)$, the baseline output is $y(1 + g_0)$, and when innovations fail, which occur with probability $1 - \lambda$, output between two consecutive periods exhibit negative growth $1 + g_0$. Basically, this can be interpreted as our model is one in which BoD's theories are attempts to improve the status quo for the firm.

In Appendix A, we show that, for given g, the present value of output is

$$Y = \frac{y}{r - \lambda g} \tag{1}$$

where r is a discount factor.

4.2 CEO Behavior and Compensation

We assume that the compensation of the CEO is a share w of the firm's output. Thus, the CEO maximizes wY by choosing z, after replacing $y = A(1-z)^{\frac{1}{2}}$ and $\lambda = z$ in (1). This yields

$$z = 2 - \frac{r_c}{g} \qquad \text{if} \quad \frac{r_c}{2} < g < r_c \tag{2a}$$

$$z = 0 \qquad \text{if} \quad 0 \leq g \leq \frac{r_c}{2} \tag{2b}$$

where r_c is the CEOs' discount rate and to keep growth bounded we assume $g < r_c$. If we replace λ and y using the optimal z in $W \equiv wY$, we obtain:

$$W = \frac{wA}{2\left[g\left(r_c - g\right)\right]^{\frac{1}{2}}} \qquad \text{if} \quad \frac{r_c}{2} < g < r_c \tag{3a}$$

$$W = \frac{wA}{r_c} \qquad \text{if} \quad 0 \leq g \leq \frac{r_c}{2} \tag{3b}$$

4.3 Shareholders returns

The share of output for the shareholders is:

$$R = \frac{(1-w)y}{r_s - \lambda g} \tag{4}$$

where r_s is the shareholders' discount rate, and to keep expected growth bounded we assume $g < r_s$. If we replace λ and y using the optimal z from 2a and 2b in (4), we obtain:

$$R = \frac{(1-w)\,A\,(r_c - g)^{\frac{1}{2}}}{(r_c + r_s - 2g)\,g^{\frac{1}{2}}} \qquad \text{if} \quad \frac{r_c}{2} < g < r_c \tag{5a}$$

$$R = \frac{(1-w)A}{r_s} \qquad \text{if} \quad 0 \leq g \leq \frac{r_c}{2} \tag{5b}$$

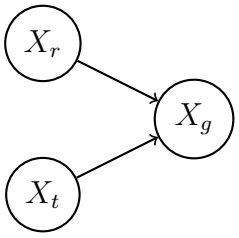
5 BoD's Theory

5.1 The Theory of Firm Growth

In our framework, the BoD's theory has two components. The first component is a theory about the future of the firm which generates an expected probability of g. The second component is about the ideal CEO profile.

Following Camuffo et al. (2023b), we represent theories through Bayesian networks (Pearl, 2009). Bayesian networks are defined by: 1) a directed acyclic graph in which each node is an attribute of the problem with uncertain realizations (a random variable); 2) a global subjective probability distribution, which can be factorized into smaller local probability distributions contingent on the edges of the graphs. The edges represent causal links among the attributes and they are the building blocks of the theory.

The Bayesian network that represents Toyota BoD's theory about the future of the firm in an "fully electrified" auto industry is:



To simplify, let attribute $X_g = \{x_g\}$ take two realizations $x_g = \{g, 0\}$ with $g \in (0, r_c)$ so that $x_g = g$ corresponds to R defined in (5a) and $x_g = 0$ corresponds to R defined in (5b).[5] The other two attributes $X_t = \{x_t\}$, and $X_r = \{x_r\}$, have continuous realizations x_i, $i = t, r$. They measure improvements in battery technology and regulatory pressures towards full electrification, which, according to this theory, raise demand and make Toyota's growth more likely. The three attributes of this problem define a space $X = X_g \times X_t \times X_r$ such that each element of X is a combination of the realizations of the three attributes, and it is a state of this space.[6]

The probability distribution of X_g is

$$X_g = \begin{cases} g & p_g \\ 0 & 1 - p_g \end{cases} \tag{6}$$

The probability p_g is itself a random variable because it depends on the realizations of the two antecedent attributes. A specific value of p_g is the marginal probability of all the states in which $X_g = g$ obtained from a probability distribution of the realizations of the three attributes defined over the space X.

Let $\pi(\cdot)$ denote probabilities. The expected value of p_g conditional on x_t, x_r and θ_g is

$$\mathbb{E}(p_g \mid x_t, x_r, \theta_g) \equiv \bar{v}_g(x_t, x_r, \theta_g) = \int_0^1 p_g \pi_g(p_g \mid x_t, x_r, \theta_g) dp_g \tag{7}$$

[5] In Appendix B we develop a full-fledged model in which x_g is a continuous measure.

[6] The particular network we choose here for illustrative purposes does not limit our model. Any type of Bayesian network would be feasible.

where θ_g is a set of parameters of this conditional probability distribution that account for the impacts of x_t and x_r on the probability of p_g. Integrating over the two antecedent attributes, obtain

$$\mathbb{E}(\bar{v}_g \mid \theta_t, \theta_r) \equiv v_g(\theta) = \int_{X_t, X_r} \bar{v}_g \pi_x(x_t, x_r \mid \theta_t, \theta_r) dx_t dx_r \tag{8}$$

where θ_t and θ_r are parameters that raise the likelihood of higher values of x_t and x_r, and $\theta = \{\theta_g, \theta_t, \theta_r\}$ is the set of all the parameters.

Different values of the parameters θ denote different probability distributions and different values of v_g.

Definition 1 (theory). *A theory is a restriction on the parameters θ.*

In our case, Toyota's theory restricts the parameters θ to values such that increases in the beliefs θ_t and θ_r raise the expected probability v_g of high growth – that is, this theory restricts the values of the parameters to the cases in which $\frac{\partial v_g}{\partial \theta_i} > 0$, $i = t, r$.[7]

Let Θ_g be the set of parameters θ that satisfy the previous restriction. We define

$$\mathop{\mathbb{E}}_{\theta \in \Theta_g} v_g \equiv \mu_{\Theta_g} = \int_{\theta \in \Theta_g} v_g \pi_{\Theta_g}(\theta) d\theta \tag{9}$$

to be the expected p_g under the restrictions of the theory, where π_{Θ_g} is a subjective probability distribution of the parameters. The expected value μ_{Θ_g} is the *expected p_g assuming that the theory is true.*[8]

[7] For example, let p_g follow a beta distribution with expected value $\bar{v}_g = \frac{e^{\theta_g x}}{1 + e^{\theta_g x}}$, where we focus for simplicity on one attribute. Let x take realizations 1, with probability θ_x, or 0. Then, $v_g = \frac{e^{\theta_g}}{1 + e^{\theta_g}} \theta_x + \frac{1}{2}(1 - \theta_x)$, which increases with θ_x *iff* $\theta_g > 0$. The theory restricts θ_g to be positive.

[8] Continuing with the example in the previous footnote, let $\theta_x = 1$ and, consistently with the theory $\theta_g > 0$, let θ_g be uniformly distributed in $(0, 1)$. Then $\mu_{\Theta_g} = \int_0^1 \frac{e^{\theta_g}}{1 + e^{\theta_g}} d\theta_g = \ln \frac{1 + e}{2}$.

Following Camuffo et al. (2023b), because decision-makers know that their theory may not be true, they set a belief $\omega_g \in (0,1)$ that the theory is true. Moreover, let $\mu_{\tilde{\Theta}_g} < \mu_{\Theta_g}$ be the expected p_g under the null hypothesis, which is a subjective expectation under alternative attributes or causal links with respect to the attributes and causal links of the theory. Decision-makers may have in mind specific alternative attributes or causal links, or they may simply provide a best guess of what is the expected probability of p_g if the theory is not true.[9] The *expected* p_g is then

$$\mu_g = \omega_g \mu_{\Theta_g} + (1 - \omega_g)\mu_{\tilde{\Theta}_g} \tag{10}$$

Definition 2 (**better theory**). *A better theory restricts to a greater extent the parameters θ, generating a higher expected probability μ_{Θ_g} under the theory.*

Better theories stem either from a stronger belief in the impact of the attributes or from the choice of different attributes that decision-makers believe have stronger effects on the states of interest. Intuitively, they are better theories because the attributes have a greater power in explaining changes in the probability of g – that is, a given change in the level of the attribute generates a more marked change in the probability of g. Since the null hypothesis is unaffected, better theories raise the expected value μ_{Θ_g} of the theory.[10]

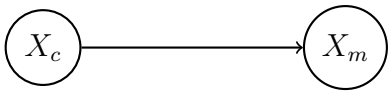
[9] A null hypothesis could be $\theta_g = 0$ with probability 1. Then, $\mu_{\tilde{\Theta}_g} = 0.5$. Theory and null hypothesis imply two different beliefs, probability distributions, and space of parameters. They are not two distributions conditional on two subspaces of the same space of parameters.

[10] In our example, suppose that the theory restricts θ_g to be uniform in $(0,2)$ rather than $(0,1)$. This is a better theory in that $\theta_g > 0$ but the probability mass has shifted towards higher θ_g. This yields $\mu_{\Theta_g} = \frac{1}{2} \ln \frac{1+e^2}{2}$, which is higher than if $\theta_g \in (0,1)$.

5.2 The Theory of CEO

Akio Toyoda also had a theory about the ideal CEO profile for his theory of the future of the firm. The profile he defined comprised several attributes such as adherence to Toyota's values, love for cars, energy, sense of pursuit, and knowledge of BEVs. Moreover, equation (5a) implies that $r_c = r_s$ maximizes the BoD's objective function R.

The attributes of the CEO-theory are CEO characteristics that the Board believes affect R in desired ways. In particular, they are factors that affect productivity A or the match of discount factors $r_c = r_s$. To streamline our analysis, we focus on one attribute $X_c = \{x_c\}$, where x_c is a set of characteristics that raises R because of a closer match of the CEO profile in terms of higher A or $r_c = r_s$. For simplicity, we do not distinguish between these two cases. However, since they are complementary, in that the marginal value of R with respect to the match $r_c = r_s$ increases with A, we can discuss them together, and make inferences on one or the other. The end-attribute is $X_m = \{m, n\}$. For simplicity, we assume that m denotes a match $r_c = r_s$ and a higher A, and n, for no-match, denotes $r_c \neq r_s$ and a lower A.[11] The Bayesian network is



Following the same steps of the previous section, replacing subscript g with subscript m, we obtain

$$\mu_m = \omega_m \mu_{\Theta_m} + (1 - \omega_m)\mu_{\tilde{\Theta}_m} \tag{11}$$

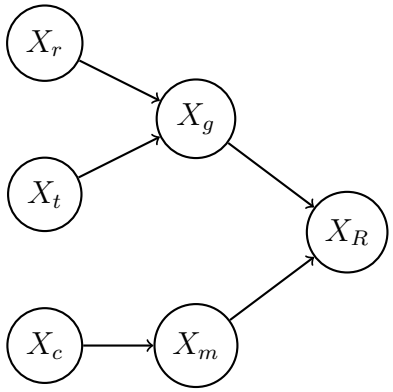
which is the expected probability generated by a theory about the possibility of finding a CEO profile associated with a higher A and $r_c = r_s$. The expected probability of match under the theory and the null hypothesis are μ_{Θ_m} and $\mu_{\tilde{\Theta}_m}$, and $\omega_m \in (0, 1)$ is the prior that the theory is true.

[11] The model in Appendix B uses a continuous attribute X_m.

6 BoDs' Choice of Theories and CEOs' Sorting

6.1 Choice of Theories

The BoD's theories of growth and CEO concur to determine performance R. The Bayesian network below combines the two theories and represents the full BoD's theory



where $X_R = \{R\}$. Using the notation R_{ij}, with $i = g, 0$ and $j = m, n$, the expected returns $\mathcal{R}$ are:

$$\begin{aligned}\mathcal{R} \equiv & R_{gm}\ \mu_g\mu_m + R_{0m}\ (1-\mu_g)\mu_m + \\ & R_{gn}\ \mu_g(1-\mu_m) + R_{0n}\ (1-\mu_g)(1-\mu_m)\end{aligned} \tag{12}$$

where R_{ij} denote R in equations (5a) or (5b) after replacing g or 0 for the growth rate, and $r_c = r_s$ and higher A or $r_c \neq r_s$ or lower A. We assume for simplicity that the determinants of g and m are independent.[12]

BoDs choose how much to invest in better theories to optimize performance. In the background there are costly efforts and endowed ability to develop better theories, which, for simplicity, we do not introduce explicitly. It suffices to say that costs ensure that there is an optimal effort to develop theories, and a greater ability to develop theories

[12] In Appendix B we remove the assumption of independence.

will lead to better optimal theories. We then focus on the marginal effects on $\mathcal{R}$ to draw implications for the behavior of the Board and performance.

More importantly, we make two remarks. The first one is the distinction between *strategies* and *theories*. Strategies could be descriptions of what decision-makers intend to do. Strategies that use theories hinge instead on the identification of attributes, causal links and probabilities, as discussed in the previous sections. The second remark is that Boards or CEOs focusing on actions take the probabilities μ_g and μ_m as given and optimize resources to generate the optimal growth rate g or to find the ideal CEO profile. In our case, decision-makers look for attributes and causal links (a theory) that maximize μ_g and μ_m, and only after they settle on these probabilities, they optimize actions. Therefore, theories add an antecedent layer to the choice of actions. This is the step of *problem formulation* (Nickerson and Zenger, 2004) in which the search for optimal theories makes firm's growth and CEO-match more likely before investing in actions or resources.

Proposition 1. *BoDs with better theories about the future of the firm (high μ_g) also have better theories of the ideal CEO (high μ_m), and vice versa*

Proof. $R_{gm} - R_{gn} - (R_{0m} - R_{0n}) > 0$ because, using (5b), $x_g = 0$ implies $R_{0m} = R_{0n} = 0$, and $R_{gm} > R_{gn}$ because, with $x_g = g \in (\frac{r_c}{2}, r_c)$, $x_m = m$ maximizes R in (5a). The cross-partial of (12) with respect to μ_g or μ_m is then positive. Thus, changes in the parameter set Θ_g that increase μ_g raise the marginal value of changes in the parameter set Θ_m that raise μ_m. As a straight application of the theory of complementarities (Milgrom and Roberts, 1990), in equilibrium BoDs that develop better theories of growth also develop better theories of the CEO match. ■

The insight of Proposition 1 is that BoDs putting more effort in developing better theories about the future of the firm will also exert more effort to develop theories about the right CEO profile for the job. Note the asymmetry. BoDs that cannot come up with

better theories of growth, will be content with generic CEO profiles. Recall that high-growth theories encourage CEOs to contribute to the development and implementation of theories instead of execution (higher z). Overall, this leads to a virtuous circle triggered by these complementarities.

The following corollary reinforces these complementarities:

Corollary. *BoDs with greater ability to envision better theories, or with a longer term horizon, develop better optimal theories of growth and better optimal theories of CEO fit.*

Proof. When all choice variables are complementary to each other, Milgrom and Roberts (1990) show that exogenous factors that raise the marginal value of one of these variables without reducing the marginal value of the others, raise the marginal value of all the choice variables. A stronger ability to develop better theories, raises the marginal value of all theories, improving all theories. Likewise, a lower r_s increases all the marginal effects of R_{ij}. Thus, a lower r_s also triggers stronger complementarities. ■

This corollary states that the complementarities have reinforcing effects. BoDs with better ability to develop theories or long-term horizons exert more effort in producing high-growth theories and more accurate and fitting CEO profiles.

6.2 Sorting CEOs

We show that CEOs with lower discount rates, hired by BoDs with matching discount rates, accept lower shares of output as compensation. These CEOs rely on their ability to contribute to a larger output in the future rather than a larger share of a smaller output in a shorter period. CEOs who ask for high compensations are more likely to be CEOs in firms with less ambitious theories of growth. In these cases, the potential CEOs cannot lever the growth of the firm to earn more output in the future, and therefore seek more

output now in the form of a higher share of an output that is not likely to grow much in the future.

Assume that there is a continuum of BoDs (companies) or CEOs ranked from low to high discount factors r_s or r_c within a given range. Since the optimal μ_g and μ_m are complementary, they move in the same direction. Therefore, BoDs higher in the ranking in terms of r_s will have lower μ_g and μ_m. We also assume for simplicity that the ideal CEO characteristics according to the BoD's theory are such that r_s and r_c closer to a match $r_s = r_c$ show a higher A. In this way, the complementarity enables us to deal with higher A and closer match as one variable.

In addition, we assume that when BoDs approach candidates, they reveal them their theory of growth and the theory of the CEO profile. As a result, CEOs know the BoDs theories and we assume that they share their beliefs μ_g and μ_m. Albeit we do not model the selection of the candidates, the sharing of theories and beliefs happen with any candidate that they approach, and BoDs make them a take-it-or-leave offer. All potential candidates then make their choice knowing the rank of the company.

Let $Y_i^S(j)$ be the evaluation of present output by the BoD. The indices $i, j = 1, 2, ...$ denote, respectively, the rankings of r_s or r_c (with lower indices denoting lower values). The superscript S denotes shareholders. This output is (12) in which the four terms R_{ij} from (5a) or (5b) are not multiplied by $1 - w$. Let $Y_i^C(j)$ be the analogous evaluation of present output by the CEO. It is equivalent to (12) with W_{ij} from (3a) and (3b) in lieu of R_{ij}, and now the terms are not multiplied by w. The superscript C denotes CEO.

Proposition 2. *There is a sorting equilibrium in which the CEO matches the BoD's discount factor, and lower discount factors reduce* w_i.

Proof. In equilibrium, agents have no incentives to deviate. The i^{th} CEO in the ranking, who matches the i^{th} BoD, has no incentive to deviate if $w_i Y_i^C(i) > w_k Y_k^C(i)$. By

symmetry, the condition for the k^{th} CEO in the ranking is $w_k Y_k^C(k) > w_i Y_i^C(k)$. Let i be a higher rank than k, that is i corresponds to lower r_s and higher μ_g and μ_m. The two conditions imply that the following inequality has to hold

$$\frac{Y_k^C(i)}{Y_i^C(i)} < \frac{w_i}{w_k} < \frac{Y_k^C(k)}{Y_i^C(k)}$$

It is not difficult to see that, given $g < r_c < 2g$, the cross-partial of (3a) for higher g and lower r_c is positive, which establishes $\frac{Y_k^C(i)}{Y_i^C(i)} < \frac{Y_k^C(k)}{Y_i^C(k)}$. Therefore, the inequality holds. Moreover, if i corresponds to a higher μ_g and μ_m, then $Y_k^C(k) < Y_i^C(k)$, which implies that in equilibrium $w_i < w_k$. ■

Proposition 2 shows that CEOs who align with high-growth BoDs are more inclined to accept lower shares of output as compensation, as they rely on the growth of output to increase their overall pay. In contrast, CEOs who align with low-growth BoDs tend to negotiate for a higher share of output as compensation.

7 Illustrative Examples

In this section, we present three multinational firms, a conglomerate, a business services company, and a financial services firm. These examples serve to contrast our theory as they provide Boards of Directors with different interests in theorizing. The first example shows the interest of the Board in short-term returns while the next two examples show the interest of the board in theorizing, development of a vision, and long-term growth. We discuss how each of these examples are represented in our model.

7.1 General Electric and Larry Culp

The BoD of General Electric (GE) appointed Larry Culp as Chairman and CEO on October 1st, 2018. At that time, the goal of GE's BoD was to reduce the large debt cumulated by Jeff Immelt in the aftermath of the 2008-2012 crisis through drastic restructuring of the poor-performing power business, reduction of corporate costs, and asset liquidation (General Electric, 2019). According to the BoD, they had to straighten the balance sheet and improve cash flows to stop the decline in stock performance.

The BoD's theory had three logical pillars. First, GE should focus on few big challenges and sets of technologies: the future of flight (aviation), precision health (healthcare), and energy transition (energy). Second, GE should be broken down into three separate companies, one per challenge. Third, GE should focus on perfect execution at the company level.

This theory was not particularly novel or original. It did not aim at moving the company to a higher growth path, but only at bringing it back to its standard levels in the short-term. The BoD also had a theory of the CEO profile: The CEO had to have a strong financial background, deep knowledge of the businesses, and a proven track record in execution and operational excellence. These attributes are common requirements in the sector. This theory led to the appointment of Jack Flannery in 2017. Flannery seemed to have the profiled characteristics, but failed to execute and was replaced one year later. The BoD still believed in their theory of the CEO and its fit with the theory of the future of the firm. They had simply chosen the wrong candidate.

Larry Culp was a member of the BoD of GE and stood out for his resume. He had served as CEO and Chairman of Danaher from 2001 to 2014, focusing on cost reduction, execution, and capital allocation optimization (strategic acquisitions and divestitures), which resulted in strong cash generation and shareholders' value. For the BoD, Larry

Culp was a very strong candidate, whose knowledge, skills and experience aligned well with the desired profile. In Thomas Horton's (GE Lead Director) words: "*Larry Culp has a proven track record in company transformation and delivering shareholder value. He is a strong leader with deep knowledge of industrials and technology, and an intense focus on execution, organization, and talent development.*"[13]

The BoD offered Culp a 3-year contract. The contract could be revised and extended and was prolonged for two more years. However, the 3-year window reflects the short-time horizon of the restructuring plan. Consistently with our model, the compensation package was attractive but focused on short-term incentives.[14] Moreover, Culp's and the BoD's views were aligned. In the appointment announcement, Culp stated: "*We will be working very hard in the coming weeks to drive superior execution, and we will move with urgency. We remain committed to strengthening the balance sheet including deleveraging.*" Culp's focus on execution reflects a low z in our model.

GE's financial performance significantly improved since Larry Culp took office in 2018, albeit they never returned to the pre-Flannery period. As predicted by our model, the short-term focus solved short-term problems, putting the company back on track, but did not produce the higher returns of theories focused on a change of pace in the growth of the company in the longer term.

7.2 Mastercard and Ajay Banga

The BoD of Mastercard (MC) appointed Ajay Banga as President and CEO on July 1, 2010. Banga succeeded Robert Selander, who had successfully led MC for 13 years during which the company became the main competitor of VISA in the US market in

[13] https://www.ge.com/news.

[14] According to some shareholders, even too attractive, as Culp managed to obtain 72 million dollars in compensation in 2020.

debit, credit, and prepaid card operations. Selander envisioned the future of MC as a key player in the emerging global industry for digital payments and a driver of the global transformation towards a cashless world.

The theory of MC's BoD – led by Selander and, later, by Richard Haythornthwaite (who became Chairman of MC's BoD in 2006) – was visionary and disruptive (high g in our framework). It logically connected the transition from paper-based forms of payment toward digital payments, the broadening of the technological solutions for digital payments (from cards to on-line, to mobile payments), and the emergence of Asia-Pacific as the highest-growth region for this market. The theory was long-term (low r_s in our framework), as also implied by the fact that it was further refined during and after the financial crisis of 2008-2012. The BoD saw the crisis as an opportunity to strengthen its position on debit cards and establish a successful model around mobile payments. To do so, it leveraged the fact that many banks were developing person-to-person payment applications that linked mobile phones to customers' accounts.

Corresponding to this innovative, long-term theory about the future of the firm, MC's BoD (namely Selander and Haythornthwaite) also had a theory of the CEO. The CEO had to have an innovation drive, the ability and willingness to further develop MC BoD's theory about the future of the firm (leading to high z, in our framework), and a long-term time horizon (low r_c). In addition, the CEO had to have a strong international flair, with an emphasis on the Asia-Pacific region as the regional markets with the highest potential for growth in innovative digital payments. Other attributes, quoted by Selander during conference calls with analysts, were a deep background in financial services to understand country-specific institutional and regulatory dynamics, and openness, thoughtfulness, and energy in management style.

This theory about the CEO drove the process of CEO succession, guided by Selander

and led by Haythornthwaite. Haythornthwaite recalls that the first meeting with Banga, one of the shortlisted candidates, was originally meant to test whether he was a good match, but "*quickly turned into a strategy session. We talked about what the company and our industry would look like two, five, ten, and twenty years into the future, and how we would develop the culture, talent, and teams to succeed*" (Haythornthwaite and Banga, 2021).

Furthermore, Haythornthwaite reports: "*We also agreed that if he was successful in the role, his tenure should last about 10 years,*" providing further evidence of the low r_s of the BoD and the low r_c of the CEO. Banga also put forward his ideas for MC's future: "*First, we would have to embrace emerging technologies and help build cutting-edge tools to make them more secure and seamless. Second, we should diversify by expanding into adjacent spaces such as cybersecurity and data analytics, adding value to the core offering. And third, we needed to address financial inclusion in a commercially viable, sustainable way: If our business model of offering choice in payments was to succeed, that choice had to be accessible to everybody.*" (Haythornthwaite and Banga, 2021).

Ajay Banga was selected as MC's Chairman and CEO and stayed in office for 11 years. During his mandate, MC delivered a 13 percent compound annual revenue growth and a cumulative total return for shareholders of 1,781% (around 16 times increase in the stock price) (Roberts and Mondalek, 2014).

7.3 JP Morgan Chase and Jamie Dimon

Jamie Dimon became Chief Executive Officer (CEO) of J.P. Morgan (JPM) Chase on January 1, 2006. He had been President of the company since the company's merger with Bank One Corporation on July 1, 2004. He succeeded William Harrison, who had become JPM's CEO in 2000 and Chairman of its BoD in 2001.

Harrison had risen through the ranks of Chemical Bank, merged with Chase Manhattan in 1995, and acquired JPM in 2000. Harrison had been one of the protagonists of the process of consolidation of the US banking industry which started in the 1990s and has continued till today. His theory about the future of the bank – and therefore that of JPM's Board – was to navigate the process of consolidation of the banking industry, becoming quickly a large player in the financial service industry. He summarized this view in an interview (Currie, 2004): "*...The biggest change, clearly – and it's a change that will continue to happen – is the consolidation that really began in earnest in the 1990s with the U.S. banks, and that was driven by the fact that size, and scale, and leadership position in your businesses would create more value. We were one of the first to consolidate and out of that came a vision and the strategic notion that size and leadership matter.*"

Harrison executed this theory between 2000 and 2003 but with unsatisfactory results in terms of profitability, stock performance, and revenue growth.

At the same time, Sandy Weill, with the help of Jamie Dimon, came up with an alternative, more innovative and articulated theory about banking (McDonald, 2009): the "universal bank." Weill's and Dimon's theory of the universal bank posited that the industry would be dominated by few large and diversified financial institutions. They would operate as "one-stop-shops" or "financial supermarkets," offering a full line-up of innovative financial products and services. Weill and Dimon's theory posited in particular that a financial supermarket had two advantages over traditional (separated) retail and investment banking: a) it stabilizes earnings because steady profits from retail banking smooth out the variability in trading; b) it increases sales, if different business units collaborate and feed one another.

The theory, totally novel in the early 1990s, was grounded on the fact that scale drives both efficiency in operations and revenue growth (fees) through customer profiling

and cross-selling of products and services. It also allows to scatter risks and stabilize revenues. According to our framework, this corresponds to a high-growth theory (high g) associated to a long-term horizon (low r_s). Based on this theory, in less than a decade, Weill and Dimon managed to build, through spectacular M&A-driven growth, Citigroup, the financial service industry leader.

Weill unexpectedly ousted Dimon out of Citigroup in 2000. However, Dimon successfully became Chairman and CEO of Bank One, which he quickly transformed into a top-performing retail bank. Here is where our story links back to Harrison. Harrison and Dimon had known each other for many years. From time to time they had informal discussions about their respective companies and trends in the financial services industry. They shared the idea that the future of banking would be based on a fully customer-centric organization, a further increase in scale, and on diverse revenue bases.

Harrison needed a CEO able and willing to revise, develop and improve JPM's theory in this direction. He articulated the CEO profile in terms of "*... balance [between] the needs for professional knowledge, business expertise, varied industry knowledge, financial expertise, and CEO-level business management experience...* ." The cornerstone of this revised theory of the universal bank became clearer to Harrison when, in November 2003, he started to discuss with Dimon the possibility and the potential merits of a business combination between JPM and Bank One. They concluded that a merger between the two companies, based on customer-centricity, scale and diversified revenues, would be beneficial to stockholders and customers and represented a potential disruption in the banking industry.

Harrison has repeatedly confirmed in public interviews that Dimon contributed to the development and implementation of JPM BoD's theory (high z). For example, he said: "*When Jamie Dimon took over in 2006, he carried it to the next level. And so*

today you have one of the best banks, if not the best bank, in the world, and they're doing great" (Rieker, 2013). Dimon developed and refined JPM's theory beyond the widespread idea of industry consolidation and of large financial institutions as conglomerates. In an interview right after his appointment as CEO, he articulated clearly his customer-centric view (Deutsch, 2006):

> *. . . I separate the idea of a conglomerate from what I call a natural set of related products, which is what we are. Most regional banks around the world do all of our businesses, so we didn't artificially put them together. The business grew up this way, and it grew up for a good reason. When you're a consumer, you walk in and you expect a certain kind of retail product set. When you're a small or middle-market business, you walk in and you expect a certain kind of financial-services product set. It's the same when you're a large corporation. And the fact is, the average regional bank provides cash-management services, private-client services, small-business services, retail, and middle market. So when we organize across these six business lines, it's really for three sets of customers— consumers, private companies, and large companies and institutions—and we provide a natural product set that we are always trying to expand. And the reason this works is that our size and scale allow us to find ways to deliver those products better, faster, and cheaper to the customer.*

Since 2006, Dimon has been at the helm of JPM outperforming competitors in the industry during and after the 2008-2012 financial crisis, generating significant growth and shareholders' value and becoming the most reputed executive in the financial service industry (Crisafulli, 2011).

8 Conclusions

This paper combines two recent streams of literature. First, it draws on Bandiera et al. (2020), which develops a typology of CEOs, building on Kotter's (2008) concept of manager vs. leader. The former prevalently carry out routine tasks and engage in internal meetings to optimize short-term performance. The latter envision the future of the firm, and spend more time articulating their vision. Bandiera et al. (2020) show that firms that hire leader CEOs perform better, which is consistent with our model. It is however unclear why not all or the majority of firms hire CEO leaders.

Second, we build on the recent literature that highlights the role of theories as a source of competitive advantage (Zenger, 2016; Felin and Zenger, 2017; Carroll and Sørensen, 2021; Camuffo et al., 2023b). Our model shows that firms that rely on theories hire leader CEOs and outperform those that do not. This directly tackles Bandiera et al.'s (2020) question about why we observe an equilibrium with few firms with leader CEOs. There are two leading explanations: Supply and Demand. It is possible that there is a limited supply of leader CEOs. We contribute to this debate with a demand-side explanation. Some firms focusing on short term (those lacking theorizing ability) may not even want leaders CEOs.

A limitation of this paper is that our theory is not easy to test. The extent to which BoDs or CEOs formulate theories can be blurred, and we need clear metrics about it. In addition, a key element of our theory is the extent to which BoDs or CEOs have short- or long-term orientations, which is again not easy to measure. Empirical evidence or tests of our theory may then require innovations in the measurement of our concepts. Text analysis through the use of AI categorization models that focus on BoD documents and the announcements of CEO profiles can help. However, we need to devise precise and innovative metrics to measure the extent to which these texts reflect causal links in

the logic of BoDs or CEOs. We also need to understand heterogeneity effects better. This paper is silent on which BoDs or CEOs are better able to develop theories. Finally, one can also think about the interdependence of multiple LFHI decisions. For instance, how firms envision the future through theories when they need to change leadership and implement a Merger for which past data is inexistent. Both the search for valid measures and tests of the theory, and a better understanding of heterogeneity and interdependence with other LFHI decisions, theoretically and empirically, are important areas of future research on this topic.

In spite of these limitations, we believe that the distinction between forward- and backward-looking strategies, and the role of theories in representing the future to take better actions, is an important determinant of firm performance. It is a relevant area of research for both our academic understanding of business processes and their practical implications.

References

Agle, B. R., Nagarajan, N. J., Sonnenfeld, J. A., and Srinivasan, D. (2006). Does CEO charisma matter? An empirical analysis of the relationships among organizational performance, environmental uncertainty, and top management team perceptions of CEO charisma. *Academy of Management Journal*, 49(1):161–174.

Baer, M., Dirks, K. T., and Nickerson, J. A. (2013). Microfoundations of strategic problem formulation. *Strategic Management Journal*, 34(2):197–214.

Bandiera, O., Prat, A., Hansen, S., and Sadun, R. (2020). CEO Behavior and Firm Performance. *Journal of Political Economy*, 128(4):1325–1369.

Berns, K. V. and Klarner, P. (2017). A review of the CEO succession literature and a future research program. *Academy of Management Perspectives*, 31(2):83–108.

Camuffo, A., Gambardella, A., and Pignataro, A. (2023a). Framing strategic decisions in the digital world. *Strategic Management Review*, 4(2):127–160.

Camuffo, A., Gambardella, A., and Pignataro, A. (2023b). Theory-driven strategic management decisions. Discussion Paper 17664v2, Center for Economic Policy Research, London.

Carroll, G. R. and Sørensen, J. B. (2021). *Making Great Strategy: Arguing for Organizational Advantage*. Columbia University Press.

Charan, R. (2005). Ending the CEO succession crisis. *Harvard business review*, 83(2):72–81.

Cheng, J. Y.-J., Groysberg, B., and Healy, P. (2020). Your CEO succession plan can't wait. *Harvard Business Review*.

Choi, J. and Levinthal, D. (2022). Wisdom in the wild: Generalization and adaptive dynamics. *Organization Science*, forthcoming.

Cragun, O. R., Nyberg, A. J., and Wright, P. M. (2016). CEO succession: what we know and where to go? *Journal of Organizational Effectiveness: People and Performance*, 3(3):222–264.

Crisafulli, P. (2011). *The House of Dimon: How JPMorgan's Jamie Dimon Rose to the Top of the Financial World*. John Wiley & Sons, Inc.

Currie, A. (2004). The big land grab. *Euromoney*, 35(418).

Deutsch, C. G. (2006). Building the global bank: An interview with jamie dimon. *MCKINSEY QUARTERLY*, (4):50 – 59.

Elfenbein, D. W. and Sterling, A. D. (2018). (When) is hiring strategic? Human capital acquisition in the age of algorithms. *Strategy Science*, 3(4):668–682.

Falato, A., Li, D., and Milbourn, T. (2015). Which skills matter in the market for CEOs? evidence from pay for CEO credentials. *Management Science*, 61(12):2845–2869.

Felin, T. and Zenger, T. R. (2017). The theory-based view: Economic actors as theorists. *Strategy Science*, 2(4):258–271.

Finkelstein, S. and Boyd, B. K. (1998). How much does the CEO matter? The role of managerial discretion in the setting of CEO compensation. *Academy of Management Journal*, 41(2):179–199.

Finkelstein, S., Hambrick, D. C., and Cannella, A. A. (2009). *Strategic leadership: Theory and research on executives, top management teams, and boards*. Oxford University Press, Oxford UK.

Fitza, M. A. (2017). How much do CEOs really matter? Reaffirming that the CEO effect is mostly due to chance. *Strategic Management Journal*, 38(3):802–811.

Galasso, A. and Simcoe, T. S. (2011). CEO overconfidence and innovation. *Management Science*, 57(8):1469–1484.

General Electric (2019). 2019 notice of annual meeting and proxy statement. Press Release.

Hambrick, D. C. and Quigley, T. J. (2014). Toward more accurate contextualization of the ceo effect on firm performance. *Strategic Management Journal*, 35(4):473–491.

Harrell, E. (2016). Succession planning: What the research says. *Harvard Business Review*, 94(12):70–74.

Haythornthwaite, R. and Banga, A. (2021). The former and current chairs of mastercard on executing a strategic ceo succession. *Harvard Business Review*, 99(2):32.

Impink, S. M., Prat, A., and Sadun, R. (2021). Communication within firms: Evidence from ceo turnovers. Working Paper 29042, National Bureau of Economic Research.

Kaplan, S. N., Klebanov, M. M., and Sorensen, M. (2012). Which CEO characteristics and abilities matter? *The Journal of Finance*, 67(3):973–1007.

Kaplan, S. N., Sørensen, M., and Zakolyukina, A. A. (2022). What is CEO overconfidence? evidence from executive assessments. *Journal of Financial Economics*, 145(2):409–425.

Karni, E. (2022). A theory-based decision model. *Journal of Economic Theory*, 201:105444.

Kotter, J. P. (2008). *Force for change: How leadership differs from management*. Simon and Schuster.

Malmendier, U. and Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of financial Economics*, 89(1):20–43.

McDonald, D. (2009). *Last Man Standing: The Ascent of Jamie Dimon and J. P. Morgan Chase*. Simon & Schuster.

Milgrom, P. and Roberts, J. (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, pages 1255–1277.

Murmann, J. P. and Schuler, B. A. (2023). Exploring the structure of internal combustion engine and battery electric vehicles: Implications for the architecture of the automotive industry. *Industrial and Corporate Change*, 32(1):129–154.

Nguyen, B. D. and Nielsen, K. M. (2014). What death can tell: Are executives paid for their contributions to firm value? *Management Science*, 60(12):2994–3010.

Nickerson, J. A. and Zenger, T. R. (2004). A knowledge-based theory of the firm—the problem-solving perspective. *Organization science*, 15(6):617–632.

Nyberg, A. J., Cragun, O. R., and Schepker, D. J. (2021). Chief executive officer succession and board decision making: Review and suggestions for advancing industrial and organizational psychology, human resources management, and organizational behavior research. *Annual Review of Organizational Psychology and Organizational Behavior*, 8:173–198.

Ortoleva, P. (2012). Modeling the change of paradigm: Non-bayesian reactions to unexpected news. *American Economic Review*, 102(6):2410–2436.

Ou, A. Y., Waldman, D. A., and Peterson, S. J. (2018). Do humble CEOs matter? An examination of CEO humility and firm outcomes. *Journal of Management*, 44(3):1147–1173.

Pearl, J. (2009). *Causality*. Cambridge University Press, Cambridge UK.

Quigley, T. J., Crossland, C., and Campbell, R. J. (2017). Shareholder perceptions of the changing impact of CEOs: Market reactions to unexpected CEO deaths, 1950–2009. *Strategic Management Journal*, 38(4):939–949.

Quigley, T. J. and Graffin, S. D. (2017). Reaffirming the CEO effect is significant and much larger than chance: A comment on Fitza (2014). *Strategic Management Journal*, 38(3):793–801.

Quigley, T. J., Wowak, A. J., and Crossland, C. (2020). Board predictive accuracy in executive selection decisions: How do initial board perceptions of CEO quality correspond with subsequent CEO career performance? *Organization Science*, 31(3):720–741.

Rieker, M. (2013). Harrison, key player in an era of consolidation, looks back. *American Banker*, 178(48).

Roberts, D. and Mondalek, A. (2014). How mastercard became a tech company. *Fortune*, 170(2):88 – 91.

Schepker, D. J., Kim, Y., Patel, P. C., Thatcher, S. M., and Campion, M. C. (2017). CEO succession, strategic change, and post-succession performance: A meta-analysis. *The Leadership Quarterly*, 28(6):701–720.

Schepker, D. J., Nyberg, A. J., Ulrich, M. D., and Wright, P. M. (2018). Planning for future leadership: Procedural rationality, formalized succession processes, and CEO influence in CEO succession planning. *Academy of Management Journal*, 61(2):523–552.

Shen, W. and Cannella Jr, A. A. (2003). Will succession planning increase shareholder wealth? Evidence from investor reactions to relay CEO successions. *Strategic management journal*, 24(2):191–198.

Waldman, D. A., Ramirez, G. G., House, R. J., and Puranam, P. (2001). Does leadership matter? CEO leadership attributes and profitability under conditions of perceived environmental uncertainty. *Academy of Management Journal*, 44(1):134–143.

Zajac, E. J. (1990). CEO selection, succession, compensation and firm performance: A theoretical integration and empirical analysis. *Strategic Management Journal*, 11(3):217–230.

Zenger, T. (2016). *Beyond competitive advantage: How to solve the puzzle of sustaining growth while creating value*. Harvard Business Review Press.

Appendix A: Long Term Output

In the first period, the CEO invests time z on contributing to develop the BoD's theory and implement it, but as discussed in the text it does not yet generate increases in output. This is consistent, for example, with Impink et al. (2021) who find that CEOs need time to develop and implement their vision once they are hired. Then, from the second period onward, output grows at rate g with probability λ in each period.

At the end of the second period, the present value of the output is

$$Y' = \frac{(1-\lambda)(y+Y') + \lambda(1+g)(y+Y')}{1+r} = \frac{(1+\lambda g)(y+Y')}{1+r}$$

where r is a discount factor and Y' is the present value of the future output stream. This is a standard Bellman equation in which at the end of the period output is $y+Y'$ if the theory that the CEOs have contributed to develop and implement has not generated productivity increases, which occurs with probability $1-\lambda$, or $(1+g)(y+Y')$ if productivity has increased, which occurs with probability λ. Solving for Y', we obtain

$$Y' = \frac{1-\lambda g}{r-\lambda g} y$$

At the beginning of the first period the present value of output is then:

$$Y = \frac{y+Y'}{1+r} = \frac{y}{r-\lambda g}$$

Appendix B: Extended Model

We extend our model to continuous realizations of g and r_c and we allow for correlations between r_c and g.

We set $X_g = \{g\}$ with $g \in (0, r_c)$, and instead of X_m we use $X_{r_c} = \{r_c\}$ with $r_c \in (0, \bar{r}_c)$. Let

$$\mathcal{R} = \int_0^{r_s} \mathcal{R}_{r_c} d\Pi_{r_c}(r_c) + \int_{r_s}^{\bar{r}_c} \mathcal{R}_{r_c} d\Pi_{r_c}(r_c)$$

where Π_{r_c} is the cumulative distribution of r_c. We distinguish between the cases in which $r_c < r_s$ and vice versa because, from (5a) in the text, R increases or decreases with respect to r_c depending on whether we are in the former or the latter case. Finally

$$\mathcal{R}_{r_c} = \int_0^{r_c} R d\Pi_g(g \mid r_c)$$

where is defined by (5a) or (5b) and $\Pi_g(g \mid r_c)$ is the cumulative distribution of g conditional on r_c.

We leave r_s and the realizations of X_t, X_r, and X_c in the background; they are in the conditioning set of the distributions. Also, in what follows we study the effects on $\mathcal{R}$ of changes in the distributions Π. These changes may come from changes in parameters that affect the distributions of the three antecedent attributes. We leave in the background that we also need to integrate $\mathcal{R}$ above with respect to the parameters of all these distributions under the theory and null hypothesis, as we discussed in the text. The effects that we

show below on $\mathcal{R}$, along with the assumption that they change $\mathcal{R}$ under the theory more than under the null hypothesis, are sufficient to make our point.

Integrating $\mathcal{R}_{r_c}$ by parts we obtain

$$\mathcal{R}_{r_c} = \int_{\frac{r_c}{2}}^{r_c} -\frac{\partial R}{\partial g}\ \Pi_g(g \mid r_c)dg$$

which uses the fact that, for $g \in (0, \frac{r_c}{2})$, R does not depend on g.

We can then integrate $\mathcal{R}$ by parts to obtain

$$\mathcal{R} = \mathcal{R}(\bar{r}_c) + \int_0^{r_s} \int_{\frac{r_c}{2}}^{r_c} \left[\frac{\partial^2 R}{\partial g \partial r_c}\ \Pi_g + \frac{\partial R}{\partial g}\ \frac{\partial \Pi_g}{\partial r_c} \right]\ \Pi_{r_c} dg dr_c$$

$$+ \int_{r_s}^{\bar{r}_c} \int_{\frac{r_c}{2}}^{r_c} \left[\frac{\partial^2 R}{\partial g \partial r_c}\ \Pi_g + \frac{\partial R}{\partial g}\ \frac{\partial \Pi_g}{\partial r_c} \right]\ \Pi_{r_c} dg dr_c$$

In this expression $\mathcal{R}(\bar{r}_c)$ is $\mathcal{R}$ evaluated at $\bar{r}_c$ and

$$\frac{\partial^2 R}{\partial g \partial r_c} = \frac{(1-w)A(r_s - r_c)}{2\left[(r_c - g)g\right]^{\frac{1}{2}}\ (r_s + r_c - 2g)^2}$$

which implies that the sign of this derivative is positive in the first integral and negative in the second integral.

If g and r_c are independent – that is, the factors that affect them are different and $\frac{\partial \Pi_g}{\partial r_c} = 0$ – factors that raise the probability of high g make the convergence of r_c towards r_s more valuable. In the first integral, the cross-partial is positive and increases in the probability of high g lower Π_g making the change in the square bracket negative. Increases in r_c make it closer to r_s. Higher probability of higher r_c lowers Π_{r_c}, which makes this change positive. In the second integral, the cross-partial is negative and we look for decreases in r_c, which raise Π_{r_c}, which establishes the complementarity.

Complementarity between growth and match in the discount factors imply that r_c lowers the probability of high growth in the first integral and increases it in the second. Since $\frac{\partial R}{\partial g} < 0$, this reinforces the complementarity above. Otherwise, in order to preserve the complementarity between the theory of growth and the theory of the CEO, the counterveiling effect of r_c and g has to be sufficiently weak.