Organizational Dynamics: Culture, Design, and Performance*

Timothy Besley
LSE and CIFAR

Torsten Persson
IIES and CIFAR

December 2017

Abstract

This paper examines the two-way interaction between organizational culture and a key aspect of organizational design, namely the choice between centralization and decentralization. We model culture via the share of managers in an organization that adopt one of two mangerial types, which affects the way they choose projects and internalize the payoffs of other managers. Using a class of cultural dynamics based on the relative payoffs of each type, we investigate the conditions under which different cultures become dominant. Our general model delivers insights into the interplay between organizational design and culture, the coexistence of different organizational cultures, the emergence of dysfunctional cultures, and organizational resistance to change. We apply special cases of this general framework to the behavior of bureaucracies, firms, and political parties.

*We are grateful to Philippe Aghion, Maitreesh Ghatak, Bob Gibbons, Oliver Hart, Jay Lee, Rocco Macchiavello, and Jean Tirole as well as participants in meetings with the IOG program of the Canadian Institute for Advanced Research, the ERINN network, the Growth and Institutions program at Tsinghua University, the 2017 PEG Conference, the 2017 ESS conference, and an IIES seminar for helpful discussions and comments, and to the European Research Council for financial support.

1 Introduction

Organizations perform very differently, even though they operate with seemingly common technologies and other conditions. Within economics, it is commonplace to argue that *strategic design* of organizations shape their performance, and a large literature studies the determinants of organizational boundaries and authority structures (Coase 1937, Williamson 1979, Grossman and Hart 1986, and Aghion and Tirole 1997). Outside economics, a very different tradition sees differential performance as reflecting different organizational *cultures* (Whyte 1956, Hofstedte 1984, Wilson 1989, and Schein 1990). Despite widespread sympathy for this idea, few agree on how to model organizational culture. Moreover, insights from the strategic-design and cultural approaches – though each puts the finger on something important – have not really been combined.

In this paper, we try to build a formal bridge between the two ways of thinking about organizations by studying the joint dynamics of organizational culture, design and performance. We first develop a general model where a changing organizational culture interacts with a changing organizational design. Then, we study four specific applications of this general model to bureaucracies, firms, and political parties.

We portray the organization's design problem as a choice between centralization and decentralization. That choice has the standard components of existing microeconomic models, such as the one in Aghion and Tirole (1997). Thus a leader who decides whether to decentralize trades off benefits, due to better use of local information, against costs, due to conflicts of interests or foregone coordination.

Our modeling of culture and its evolution deserves more discussion. We represent organizational cultures as internalized values that influence behavior and reflect the size of the group holding common values. This accords with the ideas of Edgar Schein, who defines organizational culture as

> "a set of basic tacit assumptions about how the world is and ought to be that is shared by a set of people and determines their perceptions, thoughts, feelings and, to some degree, their overt behavior." Schein (1996, page 11)

Given this approach, we suppose that culture evolves over time as it is transmitted to new members of the organization, which can alter the prevailing

social group identities among its managers. Our emphasis on social identities follows Ashforth and Mael (1989) in sociology and organizational behavior, and Akerlof and Kranton (2000) in economics. The critical assumption is that an individual derives her identity not only from the organization at large, but also from her own peer group. Identification with a particular group also means internalizing its values, such that individuals perceive a stronger affinity with group members and more likely conform with group norms. Ashforth and Mael (1989) stress, in particular, how emerging group loyalties interact with identities.

The framework we propose has two distinctive features. For given identities – a given culture – the organization's leader is more likely to decentralize project choices to a next layer of top managers when the culture is congruent with the leader's own objective. Also, expected decentralization decisions affect the cultural dynamics in a process where the share of a certain identity type depends on its relative fitness in the organization, along the lines of Boyd and Richerson (1985). Because of these two features, equilibrium organizational design and culture evolve together, each feeding back onto the other.

Our approach gives insights into a range of phenomena. First, it highlights the joint determination of organizational culture and design. Second, multiple steady-state cultures may exist for the same fundamentals, such as technology and market conditions. The performance of two organizations operating in similar environments may therefore diverge, as their cultures become entrenched. Third – and partly as a corollary of the second phenomenon – long-run dysfunctional cultures can emerge. A certain culture can thus dig in, even though it does not serve the organization's performance by standard criteria. Fourth, cultures form basins of attraction, which make organizations less responsive to shocks to their environment.

To breathe life into our canonical model, we study four specific applications. One, to public bureaucracies, stresses dilemmas of top-down control and differential performance by organizational units. Another application, to private firms, demonstrates how observed correlations between productivity, culture (management style), and organizational design can arise endogenously over time, and asks if market competition weeds out dysfunctional cultures. A third application illustrates how a strong organizational culture may become a barrier to innovation when a firm tries to adapt to new market conditions. A final application shows how two competing political parties can develop different cultures, and how one of them can be systematically more

successful, due to the interplay between party culture and effort by party workers.

The next section discusses some related research, while Section 3 brings up IBM's difficulties in the 1980s as a motivating example. Section 4 develops our canonical model of cultural dynamics and organization design, while Section 5 analyzes its static and dynamic equilibria. Section 6 applies this general model to bureaucracies, firms, and political parties. Section 7 concludes. Proofs of lemmas and propositions are collected in an Appendix.

2 Related Literature

A number of approaches try to explain why organizations might have distinctive cultures.[1] Economists use two main approaches to study corporate cultures. The most common of these focuses on beliefs. Kreps (1990) models culture as a belief-based norm in a game played by overlapping generations of agents, where cooperation is sustained by a threat of poor future performance. Tirole (1996) also adopts a generational structure to study the interplay between individual and collective reputations, which can produce persistent differences in organizational performance indicators such as levels of corruption. Similarly, Dessi (2008) emphasizes how information transmission by older generations may create distinct cultures through collective memory, while Benabou (2013) argues that organizations may develop common sets of beliefs that induce "group think." Unlike this approach, the cultural mechanism in our paper is not based on beliefs.

Another strand in the literature argues that organizations may evolve distinctive cultures through the types of people they employ. How shared motives of employees might affect their decisions has thus been explored by Besley and Ghatak (2005) and Prendergast (2007). Hodgson (1996) and Lazear (1995) propose models, which are similar to our own in stressing how different types evolve over time within an organization. Dessein and Prat (2017) study yet another source of organizational distinctiveness in the form of slow-moving organizational capital, which helps shape performance.

Foundations of cultural differences have also been explored in other contexts. Greif (1994) sees them as solutions to (different) commitment problems, and he describes "collectivist" cultures as those with beliefs more supportive of cooperation. An alternative approach taken by Akerlof (1976) and

[1] See Hermalin (2001) for an excellent review.

Akerlof and Kranton (2000) embeds manifestations of culture in preferences that drive individual behavior. Our modeling builds on the latter.

As just mentioned, microeconomic research often models corporate cultures as shaping the beliefs governing individual behavior, while ignoring the underlying values which mediate those beliefs. This approach contrasts with most treatments of culture outside of economics. For example, in their influential book on culture and organizations, Hofstede et al (2010) use the term "software of the mind" to describe the role of culture and regard underlying values as the deepest embodiment of culture.

Our approach also builds on models of cultural evolution, inspired by the research of Cavalli-Sforza and Feldman (1981), and Boyd and Richerson (1985). Studies of socialization and cultural economics has grown in recent years; Bisin and Verdier (2011) survey this field. Our specific model of cultural change through the dynamics of values – rather than dynamics of behavior or beliefs – follows the lead of Güth and Yaari (1992), Güth (1995), and Alger and Weibull (2013). There is also a close link between the formal structure of our specific model and the general literature on evolutionary dynamics in population games, as reviewed in Sandholm (2010).

Measurement of cultural differences has flourished, but largely outside of economics. For example, Hofstede (1984) began a body of research comparing organizational cultures across countries.[2] The well-known World Values Survey was developed as a means of examining international cultural differences (see Inglehart et al, 2004). Empirical studies of culture have also become extensive in economics (see Alesina et al, 2015 and Guiso et al, 2006 for overviews). While these have mostly studied individuals, they have also been applied to firms. For example, Guiso et al (2015) argue that corporate cultures encompassing integrity likely improve performance.

A large literature in business economics and sociology studies conflicts of interest inside firms, with many authors taking a starting point in Cyert and March (1963). Economists have asked how conflicting interests shape the delegation of decision-making, with key contributions by Aghion and Tirole (1997), Bolton and Farrell (1990), Alonso et al (2008), and Hart and Holmström (2010). This approach often highlights how a designer weighs the informational benefits of delegation against the value of coordination. In the same tradition, Bloom et al (2012) study empirically decentralization in firms,

[2]See Hofstede et al (2010) for a more recent survey of the extensive evidence that has been collected.

and find productivity gains from decentralization associated with greater levels of trust. Bandiera et al (2016) examine how CEOs use their scarce time, especially when it comes to their involvement in production vs. coordination. In our setting, conflicts of interest and delegation arise endogenously over time, via the interaction between an organization's external environment and its internal cultural evolution.

The idea that corporate culture is linked to firm performance is commonplace. A typical example is Wolcott and Lippitz (2007), who suggest that

> "Unless a company is blessed with the right culture – and few are – corporate entrepreneurship won't just happen. It needs to be nurtured and managed as a strategic, deliberate act." (page 82).

In this vein, our paper relates to the voluminous literature on culture in the field of organizational behavior (see e.g., Schein, 1990), which is more influenced by sociology, psychology, and anthropology than by economics. Researchers in this field have debated at length how organizational cultures are created, and many stress the role of charismatic founders (Schein, 1983). They have also touched on the perils of reforming established organizations – especially from the top down – and the emerging conflicts once cultures become established (see, e.g., Gelfand et al, 2015). Our focus on group identities builds a bridge from these ideas to a more economic perspective.

3 IBM's Design and Culture

To frame the ideas to follow, we highlight the case of IBM, which is the subject of many studies of organizational culture, including the classic work by Hofstede (1984). These case studies of organizational dynamics and culture partly reflect the strong ethos and charisma of the company's CEO Thomas J. Watson. Leading textbooks on the origins of corporate success, like Peters and Waterman (1982), also feature IBM as a prominent example, arguing that

> "(w)hat makes it live at these companies is a plethora of structural devices, systems, styles, and values, all reinforcing one another so that the companies are truly unusual in their ability to achieve extraordinary results through ordinary people..."

This quote highlights the importance of values and the need to encourage and motivate employees to serve the organization.

In the 1950s, IBM was a behemoth of mainframe computing with a dominant market share. By 1980, the company retained a 62% share of the mainframe-computer market. But its share of the *overall* computer market had declined from 60% in 1970 to 32% – partly by underperforrming in the fast-growing mini-computer market during the 1970s relative to its rivals. In 1979, this led *Business Week* to describe IBM as a "stodgy, mature company", a view corroborated by the 20% decline in IBM's stock price. To espouse the new personal-computer industry, the firm began developing the now-famed IBM PC, which prompted the quip that "IBM bringing out a personal computer would be like teaching an elephant to tap dance."

In the end, the transformation was made, but much was written about the difficulties when shifting focus from mainframes to networks and personal computing. Mills (1996) discussed this based on interviews with IBM management, emphasizing the balance between centralized and decentralized decision making.

> "IBM's top executives attempted to manage the corporation from the top, despite its great size and complexity, and in so doing exceeded their capabilities. But IBM is a closely integrated company, operates in only one industry, and has much synergy between its various businesses. It requires a high degree of central coordination and direction. It needs a judicious blend of decentralized operating management and centralized strategic direction. In the 1980s, IBM's executives failed to get the mixture right." (page 81).

Mills also blamed IBM's culture for the firm's limited response:

> "Is IBM the victim of a corporate culture that pushed the wrong type of executive to the top? Yes. IBM chief executives were too inbred, too steeped in the arrogance of success, and too certain of their own judgment in a time of challenge. IBM's culture contributed greatly to each shortcoming." (page 81)

The IBM case shares three key features with our model. First, organizational culture engenders a sense of belonging and a common interest among

groups of employees. Second, once entrenched, a culture can limit an organization's adaptability in the wake of changing priorities and market conditions. Third, a clash between top leaders and a prevailing management culture highlights the centralization question. In the model presentation to follow, we sometimes refer to IBM as a concrete example. Our application in Section 6.3 to firm innovation also deals with IBM.

4 Basic Framework

This section describes the assumptions about actors, conditions, objectives, and timing in our general model, while the next section states our main results and discusses their implications.

Key actors Consider an organization with a three-tier hierarchy: a leader, a set of senior managers, and a set of junior managers. The leader faithfully represents the organization's ultimate principal(s) – the owners of a firm, ministry (or customers) of a bureaucracy, or voters of a party – and shares their preferences. She picks a centralized or decentralized organizational design $o \in \{c, d\}$, where decentralization, $o = d$, may have benefits (better information) as well as costs (non-coordination, conflicting interests). She also carries out project choices (see below) if the organization is centralized, $o = c$.

The organization has a continuum of divisions with unit measure, $\omega \in [0, 1]$. Each division has an upper-tier (senior) manager. These managers choose projects if the organization is decentralized, $o = d$. Upper-tier managers come in two organization-specific types (see below), $\tau(\omega) \in \{0, 1\}$, with a share μ of type 0.

Each division has a lower-tier (junior) manager. These make long-term investments in effort e and acquire one of the two types. This period's junior managers become next period's senior managers. This is the only external hiring by the organization – all upper-tier managers are internally recruited.

States of the world and project choices An aggregate state $\theta \in \{0, 1\}$ captures the organization's prevailing environment. In the IBM example, state $\theta = 0, 1$ could capture market conditions favoring mainframes and PCs, respectively. We let β denote the probability of $\theta = 0$, which is iid over

time. We will say that the environment is predictable when β is close to 0 or 1, and that it is unpredictable when β is closer to $\frac{1}{2}$.

Each division ω has to make a binary project choice, $\rho(\omega, \theta) \in \{0, 1\}$. Under centralization ($o(\theta) = c$), the leader chooses a common $\rho(\theta, \omega) \in \{0, 1\}$ for all ω. Under decentralization ($o(\theta) = d$), the upper upper-tier manager in ω chooses $\rho(\omega, \theta) \in \{0, 1\}$. No contracting – state-contingent or not – can influence decentralized project choices. Think about this as reflecting non-verifiability of project choices.

The payoffs to projects (see below) depend on the realized aggregate state θ and the alignment of local projects with a local state $\sigma(\omega, \theta) \in \{0, 1\}$. For both values of θ, a share $\alpha \geq \frac{1}{2}$ of all divisions has $\sigma(\omega, \theta) = \theta$. Hence, parameter α gauges how well technology, demand, or cost is correlated across divisions.

Leader – organizational payoff A leader observes the aggregate state θ and the composition of division managers as captured by μ. But she does not observe the local states $\sigma(\omega, \theta)$ and the type of managers heading each division $\tau(\omega)$.

When making her design choice $o(\theta)$, the leader maximizes an objective, which is increasing in each of its three components:

$$\Pi(\lambda (2x-1)^2, \int \pi(|\rho(\omega, \theta) - \sigma(\omega, \theta)|, \theta) d\omega, e). \tag{1}$$

The first term $(2x-1)^2$ reflects coordination, as x is the (maximum) share of divisions that takes the same action ρ, a term which is maximized (at 1) when every ω makes the same choice. Parameter λ indexes the importance of coordination gains. This way of capturing coordination benefits is similar to that in research on the scope of the firm (Hart and Holmström, 2010) and on coordination in firms or other organizations (Bolton and Farrell, 1990, Alonso et al, 2008).

The second term summarizes how performance depends on the average, and state-dependent, adaptation of local projects to local conditions. Here, $\pi(|\rho(\omega, \theta) - \sigma(\omega, \theta)|, \theta)$ is the payoff to alignment in division ω. Throughout, we assume that

$$\pi(0, 0) - \pi(1, 0) = \pi(1, 1) - \pi(0, 1) > 0. \tag{2}$$

A local state aligned with the local project is thus always optimal in state $\theta = 0$, and never optimal in state $\theta = 1$. Referring to the IBM example,

in mainframe (PC) state 0 (1) payoffs are the highest if managers choose projects more directed to mainframes (PCs) by setting $\rho(\omega,\theta)=\sigma(\omega,\theta)=0$ ($\rho(\omega,\theta)\neq\sigma(\omega,\theta)$).

The third term is defined over aggregate effort in the organization, $e=\int e(\omega,\theta)d\omega$, integrating over efforts by lower-tier managers in all divisions ω.

A special case of the organization's payoff, used in some of the Section 6 applications, has

$$\begin{aligned}&\Pi\left(\lambda(2x-1)^2,\int\pi(|\rho(\omega,\theta)-\sigma(\omega,\theta)|,\theta)d\omega,e\right) \qquad (3)\\ =\;&\lambda(2x-1)^2\times\int\pi(|\rho(\omega,\theta)-\sigma(\omega,\theta)|,\theta)d\omega\times e,\end{aligned}$$

which we refer to as the "multiplicative case". We also sometimes assume

$$\pi(0,0)=\pi(1,1)>\pi(1,0)-\pi(0,1)=0, \qquad (4)$$

labeled the "symmetric case".

We assume that the leader chooses the design in period t to maximize (1) in that same period. As further discussed in Section 5.3, what is important is not the one-period horizon, but that the leader cannot commit to a policy rule for the future.

Upper-tier managers – types and choices Each upper-tier manager observes the local state at her division $\sigma(\omega,\theta)$, as well as θ. Upper-tier managers thus have better information than the leader, but their information advantage diminishes in alignment parameter α. If the organization is decentralized, $o(\theta)=d$, upper-tier managers choose the local projects $\rho(\omega,\theta)$.

As stated above, a share μ_t of managers identify with type 0 and the complementary share identify with type 1. Identification has two consequences. One consequence is a preference across projects: type-0 managers prefer $\rho(\omega,\theta)=\sigma(\omega,\theta)$, while type-1 managers prefer $\rho(\omega,\theta)=1-\sigma(\omega,\theta)$. In particular, the immediate payoffs from project choices are $e(\omega,\theta)u(|\rho(\omega,\theta)-\sigma(\omega,\theta)|,\tau(\omega))$, where $e(\omega,\theta)$ denotes effort by the division's lower-tier manager and

$$u(1,1)=u(0,0)=u>u(0,1)=u(1,0)=0.$$

Thus holding type $\tau=0$ or $\tau=1$ has no intrinsic advantage, as long as managers get their favorite choice.

Henceforth, we refer to μ_t as the *organization's culture*. In *state* $\theta = 0$, the leader's preferences align with those of *type* $\tau = 0$ managers, but clash with those of type $\tau = 1$ managers, and vice versa in state $\theta = 1$. In the IBM example, mainframe (PC) types do what the leader wants in the mainframe (PC) state. The latent leader-manager conflict thus varies across states, with θ, and across time periods, with μ_t. This conflict crucially influences the leader's willingness to decentralize projects to take advantage of local information, as in the (static) model of Aghion and Tirole (1997). Note that the leader cannot resolve such conflict by hiring, as we do not allow outside hiring of upper-level managers. And even if we would allow it, types are organization-specific. Below, we discuss how relaxing these assumptions might affect the analysis.

Upper-tier managers – values Another consequence of an upper-tier manager socially identifying with a certain type is that he values not only his own payoff, but also the payoffs to coworkers of the same type. Formally, the value of a manager with type $\tau(\omega)$ in ω is given by:

$$\begin{aligned} v(\tau(\omega), \omega, \theta) &= e(\omega, \theta) u(\left|\rho(\omega, \theta - \sigma(\omega, \theta)\right|, \tau(\omega)) + \\ &\quad \int e(\varpi, \theta) \xi(\tau(\varpi)) u(\left|\rho(\varpi, \theta - \sigma(\varpi, \theta)\right|, \tau(\varpi)) d\varpi, \end{aligned} \tag{5}$$

where $\tau(\varpi) \in \{0, 1\}$ is the type in division $\varpi \neq \omega$ with

$$\xi(\tau(\varpi)) = \begin{cases} \xi > 0 \text{ if } \tau(\varpi) = \tau(\omega) \\ 0 \text{ if } \tau(\varpi) \neq \tau(\omega). \end{cases}$$

These weights represent an organization-wide "esprit de corps", where parameter ξ captures the strength of social identities. The value of identifying with a type depends on the size of this group and thus directly on organizational culture μ_t. It also depends on the leader's equilibrium design choices and therefore indirectly on organizational culture.

Lower-tier managers – effort When entering the organization, each lower-tier manager chooses an effort level, $e \in [\underline{e}, \bar{e}]$.[3] As this choice is non-verifiable, no contracts can condition on effort. But effort has a private cost

[3] This effort decision is best thought of as a sunk investment which aids the productivity of the organization.

$\psi(e)$,which is increasing and convex with $\psi(\underline{e}) = 0$. The latter guarantees a minimum effort of $\underline{e}$. The payoff of lower-tier managers satisfy $e(\omega,\theta)l(|\rho(\omega,\theta-\sigma(\omega,\theta)|,\tau(\omega))$, which we interpret as a share of the upper-tier manager's "rent", with

$$l(1,1) = l(0,0) = l > l(0,1) = l(1,0) = 0.$$

We suppose that lower-tier managers decide on effort after learning the state θ, but before knowing the upper-tier manager they are matched with. If γ is the probability that $l(|\rho(\omega,\theta) - \sigma(\omega,\theta)|,\tau(\omega)) = l$, i.e., that a lower-tier manager works for a "motivated" upper-tier manager, we can then write optimal effort as

$$e^*(\gamma) = \arg\max_{e\in[\underline{e},\bar{e}]} \{\gamma l e - \psi(e)\},$$

where $e^*(\gamma)$ is increasing in γ. By the assumed timing, all lower-tier managers choose the same level of effort.

Lower-tier managers – transmission of types Cultural transmission of types/social identities from senior upper-tier managers to junior lower-tier managers is a deterministic map from μ_t to μ_{t+1}. We specify an evolutionary process, but are agnostic about the specific mechanism behind it. Specifically, we consider the class

$$\mu_{t+1} = \mu_t + q(\mu_t)Q(\Delta), \tag{6}$$

where, for all $0 \leq \mu_t \leq 1$, function Q is assumed continuous and increasing with $Q(0) = 0$ and function $q(\mu_t) \geq 0$ with $q(\mu) > 0$ for $\mu \in (0,1)$.

Argument Δ denotes *relative fitness* – i.e., the *expected* value of holding a type-0 rather than a type-1 identity:

$$\Delta(\mu) = E[v(0,\omega,\theta) - v(1,\omega,\theta);\mu], \tag{7}$$

where the expectation is taken over ω and θ. Q increasing in Δ assumes a "Darwinian" transmission process: if one type does better (in expectation), its share increases over time. However, its share remains constant when relative fitness is zero.

In the Appendix, we show that the functional form in (6) can be derived from a microfounded model where junior managers are socialized by senior managers. But it can also capture a replicator dynamic, where junior managers simply imitate more successful senior types. With the former

transmission, relative fitness depends on tomorrow's culture $\Delta(\mu_{t+1})$, with the latter it depends on today's culture $\Delta(\mu_t)$. As we will see in the next section, however, the qualitative properties of the model do not depend on this detail.

Timing The organization evolves over time, with all relevant variables indexed by t. The full timing of the model in period t is as follows:

1. The organization enters t with generation upper-tier managers, share μ_t of which has type $\tau = 0$, and the remainder has $\tau = 1$. Nature determines $\theta \in \{0, 1\}$, and $\sigma(\omega, \theta)$ for $\omega \in [0, 1]$. A new generation lower-tier managers enters
2. Lower-tier managers invest in effort $e_t \in [\underline{e}, \bar{e}]$
3. Each lower-tier manager is randomly matched with one upper-tier manager. Social identities are transmitted to the former, which gives μ_{t+1}.
4. The leader chooses organizational form $o \in \{c, d\}$
5. If $o(\theta) = c$, the leader chooses a single value $\rho(\theta) \in \{0, 1\}$, binding for all ω
6. If $o(\theta) = d$, upper-tier managers in each division choose $\rho(\omega, \theta) \in \{0, 1\}$
7. Payoffs are realized, upper-tier managers retire, and are replaced by the current lower-tier managers.

5 Analysis

In this section, we first study equilibrium choices in a given period with a fixed organizational culture – a fraction μ of type-0 managers. This allows us to map organizational culture into equilibrium organization design (and effort) as summarized in Proposition 1. Next, we study how fraction μ_t evolves through a dynamic process, which maps the outcomes under different designs into (changes of) organizational culture as summarized in Proposition 2. Finally, we draw four lessons from these two propositions.

5.1 Organization Design

How are $e(\omega, \theta)$ and $\rho(\omega, \theta)$ determined? This depends on whether the organization is centralized or not. Given the timing, e^* is independent of ω and hence we write $e(\omega, \theta) = e^*(\gamma(\theta))$.

Centralized control – stage 5 In a centralized organization, the leader chooses $\rho(\omega, \theta)$ at stage 5. These decisions follow (all Lemmas and Propositions are proven in the Appendix).

Lemma 1 *With centralization the leader picks* $\rho(\omega, \theta) = 0$ for $\theta \in \{0, 1\}$.

Given the payoff structure, the leader wishes to set $\rho(\omega, 0) = 0$ and $\rho(\omega, 1) \neq 1$. Thus a centralized organization always picks the same project for all divisions whatever the θ state. However, the interpretation of choosing project $\rho(\omega, \theta) = 0$ can be quite different in the two states. This means that $x = 1$,and there are no losses from lack of coordination.

To derive equilibrium effort, we note that $\gamma(0) = \mu\alpha + (1-\mu)(1-\alpha)$. In state $\theta = 0$ among the μ divisions with type-0 managers, a fraction α have positive payoffs for their manager, while among the $1-\mu$ divisions with type-1 managers $1-\alpha$ have positive payoffs. As a result, the ex ante probability of positive rents to lower-tier managers is $\gamma(0)$ and effort is $e^*(\gamma(0))$. Correspondingly, if $\theta = 1$ then $\gamma(1) = (1 - \gamma(0))$, and effort is $e^*(1-\gamma(0))$.

All in all, the leader's (and organization's) payoff is

$$\Pi\left(\lambda, \left[\alpha\pi(0,0) + (1-\alpha)\pi(1,0)\right], e^*(\gamma(0))\right) \qquad \text{if } \theta = 0$$

$$\Pi\left(\lambda, \left[\alpha\pi(1,1) + (1-\alpha)\pi(0,1)\right], e^*(1-\gamma(0))\right) \quad \text{if } \theta = 1.$$

Decentralized control – stage 6 Under decentralization, the μ divisions with type-0 upper-tier managers set $\rho(\omega, \theta) = \sigma(\omega, \theta)$. And the $(1-\mu)$ divisions with type-1 managers set $\rho(\omega, \theta) = 1 - \sigma(\omega, \theta)$. Here, we have

$$x = \max\left\{\mu(1-\alpha) + (1-\mu)\alpha, \mu\alpha + (1-\mu)(1-\alpha)\right\} \in [0, 1].$$

However, effort is now at its maximum, as all lower-tier managers (rationally) expect to share in the rents of upper-tier managers. That is, $\gamma(\theta) = 1$ for $\theta \in \{0, 1\}$ and effort is $e^*(1)$.

The leader's payoff is therefore

$$\Pi\left(\lambda(2x-1)^2, \left[\mu\pi\left(0,\theta\right)+\left(1-\mu\right)\pi\left(1,\theta\right)\right], e^*\left(1\right)\right).$$

Compared to centralization, the leader (and organization) always weakly loses from coordination since $x \leq 1$. She may gain or lose from aligned projects, depending on the values of θ, α and μ. But the leader gains from higher effort, exactly by how much depends on parameters μ and α.

Centralization versus decentralization – stage 4 Given these results, the leader will centralize or decentralize the organization at stage 4 depending on the values of μ and α, conditional on the realized value of θ. The optimal decisions are described in:

Proposition 1 *There exists* $\{\mu_L, \mu_H\}$ *with* $\mu_H > \mu_L$ *such that:*

1. $o\left(0\right) = d$ *if and only if*
$$\mu \geq \mu_H \geq \alpha$$
2. $o\left(1\right) = d$ *if and only if*
$$\mu \leq \mu_L \leq 1-\alpha.$$

Proposition 1 makes intuitive sense. Suppose the interests of leaders and type-0 managers are aligned, as when $\theta = 0$. A leader will decentralize if such managers make up a sufficiently large fraction of all upper-tier managers. In the IBM-example, the leader decentralizes in the mainframe state provided the share of mainframe-type managers is high enough. Conversely, the leader will only decentralize when $\theta = 1$ provided that sufficiently many managers are of type 1. Note that when $\lambda = 0$, and coordination is unimportant, $\mu_H = \alpha = 1 - \mu_L$.

5.2 Cultural Evolution

Having solved for the static equilibrium, we now turn to the dynamics of the organization's culture – its share of type-0 managers.

Candidate steady states The transmission rule in (6) yields three possible steady states, at corners $\mu = 0$ and $\mu = 1$ and an interior point where $\Delta(\mu) = 0$. The following result is applied below

Lemma 2 *Suppose there exists* $\hat{\mu} \in [0,1]$, such that $\Delta(\hat{\mu}) = 0$. *If* $\Delta(\mu)$ *is globally increasing, there are two stable steady states at* $\mu = 0$ *and* $\mu = 1$, *and the interior steady state at* $\hat{\mu}$ *is unstable.*

In the remainder of this subsection, we present a sufficient condition for equilibrium relative fitness $\Delta(\mu)$ to be globally increasing, and show that it implies divergent dynamics.

Equilibrium relative fitness – different cases Given the earlier expressions for (5) and (7), we can compute equilibrium relative fitness (of adopting type 0 rather than 1) for any realization of μ. To do so, we take expectations over ω and different realizations of θ, given μ. When doing this, we recall that $\gamma(\theta)$ depends on μ by the results in the previous subsection. We also take the implied equilibrium design choices according to Proposition 1 into account. There are three regimes to consider

In the first regime, $\theta = 0$ and $\mu \leq \mu_H$, so the leader optimally centralizes and sets $\rho = 0$. Then, relative fitness becomes

$$\delta_H(\mu) = ue^*(\gamma(0))[2\alpha - 1 + \xi(\mu + \alpha - 1)].$$

Note that this expression is increasing in μ for two reasons. On the one hand, the type-0 group grows such that a manager internalizes payoffs for a larger group. On the other hand, equilibrium effort goes up – recall that $\gamma(0)$ is increasing in μ – as managers have a larger chance of working with their preferred projects.

In the second regime, $\theta = 1$ and $\mu \geq \mu_L$ the leader centralizes, again setting $\rho = 0$. Relative fitness becomes

$$\delta_L(\mu) = ue^*(1 - \gamma(0))[1 - 2\alpha + (\mu - \alpha)\xi].$$

This expression has an ambiguous slope in μ, as effort is now going down, which may outweigh the positive group-size effect.

In the third regime of the complementary cases, there is decentralization and relative fitness is

$$\hat{\delta}(\mu) = [\xi[2\mu - 1]u]e^*(1).$$

In this case too, relative fitness is increasing in μ due to the positive group-size effect.

When is relative fitness increasing? Putting these pieces together, we can write the overall expression for the expected relative fitness of being a type-0 manager vs. a type-1 manager:

$$\Delta(\mu) = \begin{cases} \beta\hat{\delta}(\mu) + (1-\beta)\,\delta_L(\mu) & \text{if } \mu > \mu_H \\ \beta\delta_H(\mu) + (1-\beta)\,\delta_L(\mu) & \text{if } \mu \in [\mu_L, \mu_H] \\ \beta\delta_H(\mu) + (1-\beta)\,\hat{\delta}(\mu) & \text{if } \mu < \mu_L. \end{cases} \tag{8}$$

Notice that $\Delta(\mu)$ incorporates the equilibrium rule for state-contingent design choices (through its constituent δ functions), but not the actual design choice in period t. Note also that as μ varies from 0 to 1, $\Delta(\mu)$ changes both smoothly, off the design cutoffs μ_L and μ_H, and discretely, at these cutoffs. The parts that may render $\Delta(\mu)$ decreasing are the jumps at the cutoffs and the ambiguous slope of $\delta_L(\mu)$.

We now assume:

Assumption 1 $e^*(\mu+\alpha-2\mu\alpha) - (1-2\alpha)(\mu-\alpha)\frac{\partial e^*(\mu+\alpha-2\mu\alpha)}{\partial\mu} > 0$ for $\mu \geq \alpha$.

This assumption holds if effort is not too responsive over the relevant range. Then, we obtain

Lemma 3 *If Assumption 1 holds, then for all* $\{\mu, \beta\} \in [0,1] \times [0,1]$, *there exists* $\hat{\xi}$ *such that* $\Delta(\mu)$ *is globally increasing in* μ *for all values of* $\xi \geq \hat{\xi}$.

Lemma 3 implies that $\hat{\delta}(\mu_H) \geq \delta_H(\mu_H)$ and $\hat{\delta}(\mu_L) \leq \delta_L(\mu_L)$, so that $\Delta(\mu)$ takes an upward (downward) jump as we cross the two thresholds, μ_H and μ_L, from below (above). Moreover, $\Delta_\mu(\mu) > 0$ for all intermediate values μ, away from these thresholds. Hence, $\Delta(\mu)$ is globally increasing.

Lemma 3 says that if cultural identities are strong enough – in terms of the weight managers put on their co-workers' payoff – the group-size effect outweighs the negative effort effect under centralization in state $\theta = 1$. Then, we have a dynamic complementarity in the evolution of organizational culture. This implies divergent dynamics, which eventually drive organizational culture to a corner at $\mu = 0$ or $\mu = 1$.

Equilibrium cultural evolution To state our main result, we define a critical value of organizational culture, $\tilde{\mu}(\beta)$ in the intermediate region of (8), at which[4]

$$\Delta(\mu) = \beta\delta_H(\mu) + (1-\beta)\delta_L(\mu) = 0.$$

If β is close enough to $1/2$, then $\tilde{\mu}(\beta) \in [0,1]$ always exists and the dynamics of the model are described by:

Proposition 2 *Under Assumption 1 and $\xi \geq \hat{\xi}$, there are three cases*

1. *If β is close enough to 1, a type-0 culture emerges in the long run (i.e., $\lim_{t\to\infty} \mu_t = 1$) from any starting value $\mu_0 > 0$.*

2. *If β is close enough to 0, a type-1 culture emerges in the long run (i.e., $\lim_{t\to\infty} \mu_t = 0$) from any starting value $\mu_0 < 1$.*

3. *If β is such that $\tilde{\mu}(\beta) \in [\mu_L, \mu_H]$ then – if $\mu_0 > \tilde{\mu}(\beta)$, a type-0 culture emerges in the long run ($\lim_{t\to\infty} \mu_t = 1$), while if $\mu_0 < \tilde{\mu}(\beta)$ a type-1 culture emerges in the long run ($\lim_{t\to\infty} \mu_t = 0$).*

In the first two cases, the organization's long-run culture complies with the more frequent aggregate state. In Case 3, an intermediate range for β supports multiple stable steady states. However, for each and every initial condition for μ (and a specific value of β), the dynamics are still unique.

5.3 Insights from the model

We now discuss the general insights implied by Propositions 1 and 2. These concern four questions: (i) how do organizational cultures and designs interact? (ii) can different organizational cultures coexist under the same fundamentals? (iii) may dysfunctional cultures survive in the long run? and (iv) when do sticky organizational cultures lead to inertia in adapting to a changing environment?

[4]This is the value of μ at which

$$\beta e^*(\nu(\tilde{\mu}))[2\alpha - 1 + \xi(\tilde{\mu} + \alpha - 1)] +$$

$$(1-\beta)e^*(1-\nu(\tilde{\mu}))[1 - 2\alpha + (\tilde{\mu} - \alpha)\xi] = 0,$$

where $\nu(\mu) = \gamma(0)$.

(i) Organizational culture and design? How does the evolution of organizational culture, μ, interact with organizational design (centralized versus decentralized authority)? Propositions 1 and 2 say there is no deterministic relation between the two. But when β is high enough for Case 1, the organization sees a steadily increasing type-0 culture, together with decentralization in most periods (since $\theta = 0$ in most periods for high β). When β is low enough for Case 2, we instead see a trend towards a type-1 culture, and centralization most of the time. In both these cases, the organization looks predominantly peaceful with no conflict of interest between the center and the managers.

In Case 3, when β is in an intermediate range, either of these long-term outcomes can occur depending on the initial condition. But now we see the organization's design flipping back and forth. It looks conflict-ridden under centralization, which is associated with low productivity because of low effort.

(ii) Coexistence? Similar organizations can be on divergent paths, depending on their initial conditions. To be precise, suppose two organizations engage in the same activity, sharing parameters $\{\beta, \lambda, u, l, \xi\}$, and functional forms $e^*(\gamma)$ and Π. Assume also that parameter β lies in the intermediate range of Case 3 in Proposition 2, but the two organizations have different initial values μ_0 on opposite sides of the"critical juncture" for culture, namely $\tilde{\mu}(\beta)$. In the long run, we will then observe one organization with a type-0 culture and another with a type-1 culture.

This importance of initial cultures suggests that it would be interesting to extend the model with outside hiring. We conjecture that such an extension would give the result that bringing new managers is most effective at cultures close to critical juncture $\tilde{\mu}(\beta)$, assuming that social identities are portable across organizations.

While these are interesting observations, our analysis so far does not allow for interactions between different organizations. Since firms, bureaucracies, and political parties typically interact, this is an important omission. In the next section, we study different applications of the theory and some of these allow for implicit or explicit interactions between different organizations. In these cases, we ask if different organizational cultures may still coexist in the same market or the same polity. We also ask if stiffer competition between organizations tend to create homogenous cultures.

(iii) Dysfunctional cultures? To explore the possibility of dysfunctional cultures, we look at long-run payoffs. To obtain a sharper result, assume that the leader's payoff satisfies (3), as in most our applications below. Then, we have:

Proposition 3 *If the leader's per-period payoff is multiplicative, it is greater or smaller for* $\mu = 1$ *than for* $\mu = 0$ *depending on*

$$\beta\pi(0,0) \gtreqless (1-\beta)\,\pi(1,1) + \left[\frac{e^*(1-\alpha)(1-\alpha)\left[\pi(1,0)\,\beta - \pi(0,1)(1-\beta)\right]}{e^*(1) - e^*(1-\alpha)\,\alpha}\right].$$

As $\beta \to 1$, $\mu = 1$ ($\beta \to 0$, $\mu = 0$) Proposition 2 says that a long-run type-0 (type-1) culture emerges, which is indeed the best one from the leader's viewpoint. The interesting case is therefore a less predictable environment where β is close to $\frac{1}{2}$ and the steady state depends on initial conditions. Then, the organization may not converge to the culture that maximizes long-run payoffs. Indeed for β close to $\frac{1}{2}$, the gain to the leader from her preferred long-run culture can be arbitrarily large depending on how $\pi(0,0)$ compares to $\pi(1,1)$. Hence, highly dysfunctional cultures can emerge in the long run.

How important for this result is our assumption that the leader has but a one-period horizon? The important distinction is not whether the horizon is short or long, but whether the leader can commit or not. To see this, let us introduce some new notation. First, express the period-t reduced-form payoff as a function $\widetilde{\Pi}(\theta_t, \mu_t, o(\theta_t))$ of θ_t the aggregate state, μ_t culture (the single state variable), and $o(\theta_t)$ the state-dependent centralization/decentralization choice. Second, express μ_t as a reduced-form function $\mu_t = \tilde{Q}(\mu_0, \mathbf{o}_{t-1})$ of μ_0 its initial value, and $\mathbf{o}_{t-1}$ the history of state-dependent design choices up until period $t-1$, which includes all effects on cultural transmission via relative fitness values Δ in the law of motion for μ. Then we can write the expected discounted payoff at 0 as

$$W\left(\left[o\left(\theta_t\right)_{t=0}^{\infty}, \mu_0\right]\right) = \sum_{t=0}^{\infty} D^t \left[\beta\widetilde{\Pi}(0, \tilde{Q}(\mu_0, \mathbf{o}_{t-1}), o(0)) + (1-\beta)\,\widetilde{\Pi}(1, \tilde{Q}(\mu_0, \mathbf{o}_{t-1}), o(1))\right], \tag{9}$$

where $D \leq 1$ is a discount factor.

Suppose that the $t = 0$ leader could commit herself to a sequence of policy rules for every future period. The optimal decisions maximizing (9)would be:

$$o^*(\theta_t, \mu_0) \in \arg \max_{o(\theta_t) \in \{c,d\}} \left\{ W\left(\left[o\left(\theta_t\right)_{t=0}^{\infty}, \mu_0 \right] \right) \right\}. \tag{10}$$

This sequence could well differ from the equilibrium we have studied. In particular, a leader who starts out with a dysfunctional culture, say $\mu_0 = 1$ with $\beta < \frac{1}{2}$, may want to commit herself to a sequence of state-independent centralization, $o\left(\theta_t\right) = c$ for any θ_t to initiate a transition towards a type-1 culture. Any short-run losses will be dominated by long-run gains for inefficient enough an initial culture and D close enough to 1. The key strategic consideration is that committing to future policy rules will shape future relative fitness values Δ_t and hence future cultures.

Suppose now the leader cannot commit to future policy rules, but still maximizes (9). Operating under such discretion, she takes all future leader design choices as given. Moreover, as noted in Section 5.2, her current choice of o_t does *not* affect expected relative fitness (whether given by Δ_t or Δ_{t+1}) among current junior managers. Since these managers have a one-period horizon, the choice of o_t does not affect cultural transmission. As the leader cannot influence future cultures (state variables), there is no strategic effect on cultural dynamics to consider. Her optimal design thus simply maximizes the current payoff – i.e., the equilibrium is the one we have already studied, even if leaders have an infinite horizon.[5]

To summarize, dysfunctional cultures may emerge, not because leaders are myopic but because they lack commitment. This observation ties our model to earlier discussions around the Coase Theorem. In particular, it parallels Acemoglu (2003) who shows that lack of commitment by current decision-makers is the key impediment to efficiency in dynamic political models.

A possible substitute for commitment in our setting would be for the principals to delegate long-run control of the organization to a leader who favors one particular culture over another. This would be particularly relevant where the (unachievable) commitment path would prescribe either $o_t = c$ or

[5]Short horizons among the managers do play a role, however. If each generation of managers were to internalize the payoffs of group members, not only in their own generation but also in future generations of managers, strategic concerns among leaders may reappear.

$o_t = d$ for all time.[6]

(iv) Organizational inertia? Another upshot from the model is that culture can be immune to change, even if some parameter values are permanently altered. Organizational cultures can thus limit adaptability, as in the IBM-example discussed in Section 3 above and further in Section 6.3 below. To illustrate, consider two values β_L, β_H such that

$$\beta_H e^*(1-\alpha)[2\alpha - 1 + \xi(\alpha - 1)] - (1-\beta_H)\xi e^*(1) > 0.$$

Under this assumption, the organization will converge globally to $\mu = 1$ when $\beta = \beta_H$ (as per the condition in (8) and Proposition 2.)

What happens in such a corner solution if β suddenly shifts to β_L? Given a starting point of $\mu = 1$, we obtain a kind of hysteresis. From (8) and Proposition 2, for all β such that

$$\beta\xi e^*(1) + (1-\beta)\, e^*(1-\alpha)\,[1 - 2\alpha + (1-\alpha)\xi] > 0 \tag{11}$$

culture persists at point $\mu = 1$.[7] This is because $\Delta(1) > 0$. From (11), there exists a critical value of β, given by

$$\hat{\beta}_L = \frac{e^*(1-\alpha)\,[2\alpha - 1 - (1-\alpha)\xi]}{\xi e^*(1) + e^*(1-\alpha)\,[2\alpha - 1 - (1-\alpha)\xi]}, \tag{12}$$

below which the culture will begin to change as $\Delta(1) < 0$ for all $\beta < \hat{\beta}_L$.

This result says that only large enough shifts in the environment will initiate cultural change. As $\hat{\beta}_L$ is decreasing in ξ, the loyalty bond entailed in organizational identity, the cultural friction is greater the stronger these bonds. Our model thus conforms to frequent claims in the organizational literature that identity-based cultures naturally prevent organizational adaptability.

[6] This logic is reminiscent of that in Vickers (1985), where an oligopolistic firm seeking to maximize profits can raise profits by appointing a CEO with an objective to maximize sales as a way of committing to aggressive pricing behavior. Rogoff (1985) studies strategic delegation in the context of central banking.

[7] This implies that

$$\beta\hat{\delta}(1) + (1-\beta)\delta_L(1) > 0.$$

6 Applications

In this section, we put our general model to work. In four specific applications, we show how it can illuminate questions around the roles for culture and design of organizations such as bureaucracies, firms, and political parties. Making our general approach more specific also raises new issues, which may merit further analysis in future research.

6.1 Performance of Public Bureaucracies

One of the biggest puzzles about public organizations is the wide range of performance among units of government with similar technology and access to resources. Classic accounts of public bureaucracy, like Wilson (1989), emphasize culture and values as elements that can explain inertia and resistance to change. Because traditional performance management may have limited force in bureaucracies such as police forces, hospitals and schools – where public-service outputs are hard to measure, making incentive contracts hard to implement – good service delivery may have to rely on intrinsic motivations of detectives, physicians, or teachers. Wilson (1989) also stresses that we can think about effective bureaucracies as mission-oriented organizations employing motivated agents, a suggestion picked up by Tirole (1994) and Besley and Ghatak (2005).

Examples Applying insights from their analysis of private firms, Bloom et al (2014, 2015) find the same differences in bureaucratic management as in private management, and management styles systematically correlated with bureaucratic performance indicators. Appeals to organizational culture are commonplace in consulting reports on performance. A case in point is CHKS (2012) – a report by the leading provider of healthcare intelligence in the UK – which concluded that

> "top-performing acute sector organizations invest considerable time and effort into developing an organizational culture around the delivery of high-quality, safe and efficient care" (p. 13).

Another salient example is a university with multiple priorities, including good teaching and successful research. Corresponding to the leader in our model, a dean who internalizes university priorities may look at future fund

raising or tuition fees. However, faculty members may have their own priorities over teaching or research with a higher productivity when performing the task they value the most. Moreover, cultural transmission from senior faculty may be key to how junior faculty build such values.

Applying the model to bureaucracy In any kind of bureaucracy, it is central how much mission choice to centralize and how much local discretion to allow. Leaders may be concerned that decentralization lead the organization astray from top priorities. Our framework helps understand the challenges of building an organizational culture, which serves the ultimate beneficiaries such as crime victims, patients, or students.

We thus interpret ω as reflecting different providers in a system of police precincts, hospitals, or schools. The choices $\rho(\omega,\theta)$ reflect aspects of the mission: towards which crimes to orient resources, which medical treatments to prioritize, or what teaching curricula to develop. The variable $\theta \in \{0,1\}$ reflects the leader's beliefs about the organization's priorities, while $\sigma(\omega,\theta)$ allows local variation in the mission. Lower-tier managers are the professionals who deliver services and from whom senior management is drawn. In practice, not every front-line professional becomes a senior manager, but professionals are a major source of recruitment – school principals are often former teachers. These managers are more motivated when they undertake an activity they like.

When applying our general framework to bureaucracy, we assume the organizational objective to be multiplicative and symmetric, as in (3) and (4)

$$\Pi\left(\lambda(2x-1)^2, \int_0^1 \pi\left(\left|\rho(\omega,\theta)-\sigma(\omega,\theta)\right|,\theta\right)d\omega, e\right) = \tag{13}$$
$$\hat{\phi}(x) \times e \times \int \pi\left(\left|\rho(\omega,\theta)-\sigma(\omega,\theta)\right|,\theta\right)d\omega,$$

where $\hat{\phi}(x) = \frac{1+\lambda[2x-1]^2}{1+\lambda}$ represents possible spillovers across service providers from coordination and $\pi(1,1) = \pi(0,0) = \pi_H > \pi(0,1) = \pi(1,0) = \pi_L$. The latter assumption says there is no intrinsic advantage to any possible priority. It implies

$$\int \pi\left(\left|\rho(\omega,\theta)-\sigma(\omega,\theta)\right|,\theta\right)d\omega = \pi_L + (\pi_H - \pi_L)\left[\theta + \mu - 2\mu\theta\right]$$

under decentralization. All the results in Propositions 1-3 apply straightforwardly in this case.[8]

This model can cast light on three frequently discussed features of public bureaucracies: (i) dilemmas of top-down control, (ii) heterogeneous performance, not explained by resources or technologies, and (iii) institutional inertia and resistance to reform.

Dilemmas of top-down control How much local control to offer in the delivery of public services has been discussed in research on education and health-care provision (see e.g., Wilson 1989 and Ahmad et al. 2005). It is frequently claimed that decentralization works best to take advantage of local conditions when the objectives of the center and delivery units are strongly aligned.

However, our model does not take alignment as given, and instead emphasizes that it will evolve dynamically and reflect experience with central and decentralized control. Proposition 2 shows that culture supports the center's long-run objectives when goals are clearly defined, i.e., β is close to 0 or 1. Tension is more likely when the environment is more uncertain, so that β is close to $\frac{1}{2}$ and different cultures may emerge.

Moreover, our model suggests we should see top-down control when the center and delivery units are poorly aligned. Specifically, Proposition 2 shows that if β is close to 0 or 1, such clashing interests are unusual, which raises organizational efficiency. However, when β is close to $\frac{1}{2}$ conflict between leaders and management and inefficiency due to lower managerial effort are more common. So mission clarity is associated with better performance, as in Wilson (1989).

Heterogeneous performance Our model speaks directly to the central puzzle that bureaucratic performance may differ in apparently similar organizations. This is true in case 3 of Proposition 2, where close to $\tilde{\mu}(\beta)$ organizational units may follow different paths. If state θ is common across organizations, then at a point in time when $\theta = 0$ ($\theta = 1$) organizations with

[8] In this case, the condition in Proposition 3 boils down to

$$(1-2\beta)\left[\frac{e^*(1)-\alpha e^*(1-\alpha)}{(1-\alpha)\,e^*(1-\alpha)}\right] \lessgtr (1-2\beta)\,\frac{\pi_L}{\pi_H}.$$

a culture approaching $\mu = 1$ will perform better (worse) than those with μ approaching 0.

Institutional inertia and resistance to reform The difficulty in reforming public bureaucracies due to entrenched culture is frequently discussed in the management literature (e.g., Gioia and Thomas, 1996, for academia). To understand this in our model, imagine that parameter β permanently changes at a time where a bureaucratic organization has achieved a steady state with either $\mu = 1$ or $\mu = 0$. Then, organizational culture may not adapt at all due to an entrenched managerial culture. The organization can try to handle this by centralizing, but will suffer from low efficiency due to low effort $e^* (1 - \alpha)$ under centralization, rather than $e^*(1)$ under decentralization.

Even if the change in β is sufficiently large to set in motion a cultural dynamic towards a new steady state, this may be a slow process with the length of the transition being dependent on the generational structure of managers. It will also depend on the rate of labor-market turnover, an aspect we have abstracted from. In future work, it will be interesting to consider the role of hiring and firing and to allow for (at least partial) portability of cultures across organizations.

6.2 Firms, Productivity, and Corporate Cultures

In many ways, a public bureaucracy of civil servants is similar to a private bureaucracy of managers. The insights from the previous subsection thus largely carry over. However, a private firm may be subject to a harder budget constraint, as it has to survive in the market. To consider these issues, we apply a version of our model that can generate heterogenous firm productivities. and possibly link them to different management styles, as studied by Bloom and van Reenen with different coauthors. Specifically, we use a "span-of-control" model as in Lucas (1978), where managers in each division can hire workers and the leader is a profit-motivated CEO.

Technology Suppose the productivity of each division in the firm is given by

$$\nu \left(\left| \rho(\omega, \theta) - \sigma(\omega, \theta) \right|, \theta, e, x \right)^{1-\zeta} = \left[\hat{\phi} \left(x \right) \pi(\left| \rho \left(\omega, \theta \right) - \sigma \left(\omega, \theta \right) \right|, \theta) e \right]^{1-\zeta},$$

where $\hat{\phi}(x) = \frac{1+\lambda[2x-1]^2}{1+\lambda}$ reflects the value of coordination for productivity.[9] We maintain the symmetric case where $\pi(|\rho(\omega,\theta) - \sigma(\omega,\theta)|, \theta)$ satisfies (4). Independently of the firm's organization, the division can hire labor $l(\omega)$ with a decreasing-returns production function: $\nu^{1-\zeta} l^{\zeta}$ where $\zeta < 1$. Laborers l can be freely hired at wage w.

We can now ask how organizational culture shapes the firm's management style – embodied in $\rho(\omega,\theta)$ – which, in turn, shapes organization design. The latter choice can affect the firm's profitability, which also depends on culture as embodied in μ. Aggregate shock θ reflects different states of the world, where different management activities are more or less productive. Parameter β captures how the firm's CEO evaluates these managerial decisions. A culture clash arises when upper-tier managers have a proclivity towards activities which are counterproductive for the firm.

Hiring and profits Suppose the firm's output has price p. Then the profitability of a division optimizing its hiring decision is:

$$\max_{l} \left\{ p\nu \left(\rho(\omega,\theta), \theta, e, x\right)^{1-\zeta} l^{\zeta} - wl \right\} =$$

$$(1-\zeta)\,\hat{\zeta}(w)\, p^{\frac{1}{1-\zeta}}\, \hat{\phi}(x)\, \pi(|\rho(\omega,\theta) - \sigma(\omega,\theta)|, \theta) e,$$

where $\hat{\zeta}(w) = \left(\frac{w}{\zeta}\right)^{-\frac{\zeta}{1-\zeta}}$. In this setting, division-level and firm-level heterogeneities depend on recruitment and project decisions by upper-tier managers – think about the latter as the firm's "management style". In this sense, the model in this subsection provides a microfoundation for the empirical analysis in Bloom and Van Reenen (2007).

Firm profits – the CEO's objective – has the form in (3), i.e.,

$$\begin{aligned} & \Pi\left(\lambda(2x-1)^2, \left[\int_0^1 \pi(|\rho(\omega,\theta) - \sigma(\omega,\theta)|, \theta) d\omega\right], e\right) \qquad (14) \\ = \;& (1-\zeta)\,\hat{\zeta}(w)\, p^{\frac{1}{1-\zeta}}\, \hat{\phi}(x) \left[\pi_L + (\pi_H - \pi_L)\left[\theta + \mu - 2\mu\theta\right]\right] e. \end{aligned}$$

Profits are greater when managers put in more effort (e is high), when the firm is better coordinated (μ close to zero or one), and when divisions are better aligned with local conditions ($\rho(\omega,\theta)$ and $\sigma(\omega,\theta)$ fit better together) given state θ.

[9] We normalize by $(1+\lambda)$ so that coordinated firms do not become unboundedly more productive as λ gets large.

Centralized control, management form, and firm heterogeneity This application fits our general model, so Propositions 1-3 all apply. It therefore gives a possible foundation for Bloom et al (2012), who find that decentralized firms have better performance. However, our model predicts that decentralization, management culture, managerial effort, and firm performance are all jointly determined. Thus a complex web of causal interdependencies run between these outcomes. We should expect decentralization when this is likely to have a positive impact on performance. The model can also explain a clash between leaders who represent shareholder interests and operational managers, the former wishing to limit the discretion of the latter. This is a feature of the IBM example discussed in Section 3.

More generally, our model can explain persistent heterogeneities in productivity and profits among firms, when the same market conditions and technologies are available to them. Firms that evolve better cultures will be more productive and profitable. Our framework suggests that homogenous firms are only likely to emerge when β is close to zero or one – i.e., when the environment is highly predictable and supports one specific organizational culture. When firms may face different challenges, different cultures can emerge and one of these can be better for (average) productivity.

Market selection and inefficient cultures As mentioned, a key difference between public services and private firms is that market discipline can bound cultural inefficiencies of the latter. We now explore this idea, focusing on the case where $\lambda = 0$ – i.e., we abstract from coordination gains. To stay in business in each period, a firm has to incur a fixed cost F (in terms of labor), which is paid before θ is realized. This way, market selection may only allow firms with certain cultures to carry on operating.

Suppose that prices and wages, p and w, are exogenously fixed and that

$$(1-\zeta)\,\hat{\zeta}\,(w)\,p^{\frac{1}{1-\zeta}}\pi_H e^*(1) - wF > 0,$$

which says that a maximally efficient firm is viable given the fixed cost F. In our model, this level of efficiency is never attainable if $\beta \in (0,1)$. With an interior value of β, firms will converge to a culture which entails an efficiency loss in either state $\theta = 0$ or state $\theta = 1$, as managers have to act against their preferences.

Can both type-0 and type-1 cultures coexist, or does the market constraint make one of them infeasible? To probe this, suppose β belongs to the

range in Proposition 2, where firms may evolve into either culture $\mu = 1$ or culture $\mu = 0$.

We want to give a condition for the coexistence of both cultures. Define bounds

$$\hat{\pi}_0 = \beta \pi_H e^* (1) + (1 - \beta) [\alpha \pi_H + (1 - \alpha) \pi_L] e^* (1 - \alpha)$$

and

$$\hat{\pi}_1 = (1 - \beta) \pi_H e^* (1) + \beta [\alpha \pi_H + (1 - \alpha) \pi_L] e^* (1 - \alpha)$$

for cultures $\mu = 1$ and $\mu = 0$ respectively. Given the symmetric payoffs, $\hat{\pi}_0 > \hat{\pi}_1$ if and only if $\beta > 1/2$. Then, we have

Proposition 4 *In the long-run, cultures* $\mu = 1$ and $\mu = 0$ *can coexist iff*

$$\min \{\hat{\pi}_1, \hat{\pi}_0\} \geq \frac{wF}{(1 - \zeta) \hat{\zeta} (w) p^{\frac{1}{1-\zeta}}}.$$

The proposition bounds the inefficiency among firms with different long-run cultures. Via the LHS of the inequality, this bound depends on the predictability of the aggregate environment, β, the correlation across divisions, α, and the efficiency loss due to low effort $e^* (1) - e^* (1 - \alpha)$. Via the RHS of the inequality, the bound also depends on w, p, and F. Coexistence is more likely when fixed costs are low or real product wages are low (so profits are high), both contributing to a weak market test.

With coexistence, one culture becomes relatively dysfunctional. Which one depends on whether $\beta \gtreqless \frac{1}{2}$. Thus our model offers a particular take on the observation that firms in the same market sometimes operate with persistently different productivities. Moreover, as in Bloom and Van Reenen (2007), this could be associated with persistently different management styles with management focusing on different problems and tackling them in different ways.

If the inequality in Proposition 4 fails, the market will eventually weed out one of the cultures. Unsurprisingly, a hard budget constraint reduces long-run permissible cultural inefficiencies. Shifts in market conditions which lower p or raise w – like deregulation or opening to trade – may thus help eliminate inefficient cultures.[10]

[10] Our model also predicts that the aggregate distribution of corporate cultures will affect

6.3 Culture and Management in IBM

Section 3 brought up IBM's challenge to adapt its culture to a new product line. Our model can be used to revisit this case study. Suppose the firm can specialize in one of two products: mainframes, M and PCs, P. Let $\pi_M(\theta)$ and $\pi_P(\theta)$ be divisional profits associated with two conditions, summarized by aggregate state θ, with $\pi_M(0) > \pi_P(0)$ and $\pi_M(1) < \pi_P(1)$.[11] Finally, a share μ of managers have $\tau(\omega) = 0$ and adopt a mainframe-oriented culture, while those with $\tau(\omega) = 1$ adopt a PC-oriented culture. Hence, managers focus on projects enhancing the products they identify with.

With the multiplicative performance function in (3), the firm's profits (leader's payoff) become

$$\lambda(2x-1)^2 \left[\pi_M(\theta)\, y(\theta) + \pi_P(\theta)\,(1 - y(\theta))\right] e,$$

where $y(\theta)$ is the share of divisions that adopt mainframe-enhancing activities in state θ. Under these assumptions, Propositions 1-3 apply.

Consider a firm like the old IBM, where $\mu = \beta = 1$, due to cultural convergence as in Proposition 2. Because the state is always $\theta = 0$, this firm is decentralized, and all lower-tier managers are motivated to put in effort $e^*(1)$. Moreover, the uniform culture and work habits are fully coordinated on mainframes with $y(0) = x = 1$. Profits are thus $\lambda[\pi_M(0)]e^*(1)$.

Changing market conditions What happens if β falls, making state $\theta = 1$ more common, as PCs becomes more attractive? In state $\theta = 1$, the leader optimally responds by centralizing and imposing PC-oriented projects on all divisions, since $\pi_M(1) < [\pi_P(1)\,\alpha + (1-\alpha)\,\pi_M(1)]$. As local information is

---

the equilibrium price in a market, with more efficient cultures leading to lower market prices and hence tightening the selection condition. Suppose that there is a continuum of firms in an industry and a constant elasticity demand curve, $p = Q^{-\varepsilon}$, with elasticity ε with with $Q(\theta)$ denoting total industry output in state θ with θ common to all firms. Suppose a proportion $\Omega(\theta)$ of firms has evolved a culture where the management is aligned with the firm in state θ. Then the equilibrium price in state θ is

$$p(\theta) = \left(\hat{\zeta}(w)\left[\Omega(\theta)\,\pi_H e^*(1) + (1 - \Omega(\theta))\left[\alpha\pi_H + (1-\alpha)\,\pi_L\right]\underline{e}\right]\right)^{-\frac{\varepsilon}{1-\zeta}}.$$

Note that prices are lower in states of the world that favor the dominant industry culture.

[11] Here, $\pi_P(1) = \pi(1,1), \pi_M(1) = \pi(0,1), \pi(0) = \pi(1,0)$, and $\pi_M(0) = \pi(0,0)$

lost, this will lead to some advances in PCs and some in mainframes (by "misdirected" managers). Profits are now $\lambda\left[\pi_P\left(1\right)\alpha+\left(1-\alpha\right)\pi_M\left(1\right)\right]e^*\left(1-\alpha\right)$.

These profits are lower than the profits of a firm with a PC culture, $\mu=0$. Such a firm elicits full effort $e^*\left(1\right)$ from its managers, and can decentralize projects to get better aligned decisions. This gives profits $\pi_P\left(1\right)$ for all divisions, and aggregate profits $\lambda[\pi_P\left(1\right)]e^*\left(1\right)$. Due to lower efforts and some misdirected managers, IBM will look like "an elephant learning to tap dance", compared to firms with PC-oriented cultures.

Adaptation or not So will IBM adapt? This depends on how managers perceive the change in β. Following the analysis in Section 5, if the "death of the mainframe" is still in doubt – such that β is higher than $\hat{\beta}_L$ defined in (12) – culture may not change. This is especially likely with a strong *esprit de corps* among the managers (high ξ). If and when β falls further in the new environment, cultural change begins. But during the transition, IBM has to wait for sufficiently many managers to turn over in the socialization process.

This analysis illustrates not only the narrative of IBM and its slow adaptability due to a strong culture. It also captures similar concerns nowadays expressed about Google, as it tries to adapt to greater competition and new product lines – e.g., taking on Facebook and providing mobile apps.

This discussion suggests a trade-off. Strong organizational cultures can be very powerful in stable environments. But they create inertia and risk becoming dysfunctional when adaptation is necessary. It would be interesting in further work to combine this insight with the analysis of market selection in Subsection 6.2. We conjecture that the market may eventually weed out "dinosaur" cultures, but this process may be slower if competition is weak.

6.4 Political Parties

Finally, we apply our framework to political parties and electoral competition. Thus we consider the emergence of party cultures and their interaction with party organization, with more or less say by "mid-level" politicians. This is an understudied dimension of politics, though standard political-science treatments of parties do point out that centralized authority is sometimes needed but can become too strong (Cox and McCubbins, 2003). Some authors also point out that the organization of parties is important. For example, Willis et al (1999) argue convincingly that the differential structure of

Latin American parties – e.g., very centralized in Mexico and decentralized in Brazil – are key to understanding political powers on the continent.

Voter preferences Consider a set-up with two parties $P = A, B$. Each has a leader who runs a multi-division organization – with local (district or group) party heads and party workers as upper-tier and lower-tier managers – like in Sections 4 and 5. In each t, the leader maximizes the party's probability of winning an election at the end of the period.

Voters are partitioned into a continuum of districts, or groups, indexed by ω. All voters in district (or group) ω have identical preferences:

$$W(\theta, \omega, x, e) = \lambda (2x - 1)^2 + \pi \left(\left| \rho (\omega, \theta) - \sigma (\omega, \theta) \right|, \theta \right) \cdot e(\omega) + D_B \chi. \quad (15)$$

The first term represents some need for a nationally coordinated policy, where λ indexes the importance of coordination. The second term captures a policy targeted to district ω, magnified by the effort $e(\omega)$ local party workers put into policy design. Furthermore, voters get an extra χ of utility under party-B rule: χ being a popularity shock in favor of party B, continuously distributed with mean zero, $E(\chi) = 0$, and a symmetric single-peaked density. By symmetry, c.d.f. Π of χ has $\Pi(0) = 1/2$. The χ-shock is realized after policy-design choices at stage 5 or 6, but before the election in each period. When $\theta = 1$ ($\theta = 0$) voter preferences accord with those of type-1 (type-0) district leaders, which occurs with probability $1 - \beta$ (probability β). Again we work with (4) but normalize $\pi_H = 1$ and $\pi_L = 0$.

Winning probabilities As parties offer policies $\left\{ \rho^P (\omega, \theta), x_P, e_P(\omega) \right\}$, voters in district ω vote for party A if

$$\chi \leq W^A(\theta, \omega, x, e) - W^B(\theta, \omega, x, e).$$

Observe also that

$$\int W^P(\theta, \omega, x, e) d\omega = \lambda (2x_P - 1)^2 + e_P \cdot [(1 - \theta) x_P + \theta (1 - x_P)]$$

is a function only of aggregate choices and effort. Standard arguments allow us to write party A's probability of winning the entire election as

$$\begin{aligned} p(\theta) &= \text{Prob}[\chi \leq \int W^A(\theta, \omega, x, e) d\omega - \int W^B(\theta, \omega, x, e) d\omega] \\ &= \Pi \left(W^A(\theta, x, e) - W^B(\theta, x, e) \right). \end{aligned} \quad (16)$$

Party B's probability of winning is just given by $1 - p(\theta)$.

Substituting from (15) into (16), the probability of winning for party A is

$$\Pi\left(\lambda(2x-1)^2, e \cdot \int \pi\left(\left|\rho(\omega,\theta) - \sigma(\omega,\theta)\right|, \theta\right) d\omega - W^B(\theta, x, e)\right). \quad (17)$$

Depending on θ, the leader organizes the party (and picks a set of local policies under centralization) so as to maximize the probability of winning, taking the organization and policies of party B as given. This objective fits our general model so Propositions 1-2 apply.

Decentralization and party cultures This model hints at a novel aspect of electoral competition, which has not received much attention in the academic literature. Green parties in European countries, like Germany and Sweden, started out in the 1970s and 1980s as very decentralized organizations with a party culture where local party workers engaged in local environmental projects and resisted nuclear power. In the early 1990s, issues like German integration and the Swedish economic crisis became much more salient – this is like a shift in state θ. Moreover, Green parties came to take part in regional and national coalition governments – this is like an upward shift in λ (the weight on coordinated policies). Both shifts made party leaders adopt more centralized policies, which met complaints among party members and former party leaders. Our model would portray the changing party strategies as rational responses to a changing environment as perceived by party leaders.

The earlier analysis suggests how changing party objectives might gradually change prevailing party cultures. Following the logic of Proposition 2, different party cultures can emerge. In particular, consider a value of β in the intermediate range identified in Proposition 2, such that case 3 applies. Further, assume that the initial values of μ in the two parties lie on opposite sides of critical value $\tilde{\mu}(\beta)$. To fix ideas, suppose that

$$\mu_0^B < \tilde{\mu}(\beta) < \mu_0^A.$$

Then, Proposition 2 says that party A evolves a long-run party culture with $\mu^A = 1$, and party B one where $\mu^B = 0$, where both cultures are associated with loyal party workers. Both party cultures may coexist and, as we see below, one party could spend more time in office even if party fundamentals

are similar, on the back of a different party organization. In the long run, parties may or may not be decentralized, depending on the value of λ, i.e., how valuable coordination is to winning. Studying this further in specific party contexts would be interesting.

A competitive cultural advantage? Let's see how a party culture can become an electoral asset or liability. Consider the case where $\mu^A = 1$ and $\mu^B = 0$. The win probability for party A is then $p^A = \Pi(W^{A*} - W^{B*})$, where W^{P*} denotes the equilibrium utility offered by party P to the aggregate of voters. Party A has an electoral advantage with $p^A \gtreqless 1/2$ as $\Pi(W^{A*} - W^{B*}) \gtreqless 1/2$. Under these conditions, we have

Proposition 5 *Suppose that* $\mu_L > 0$ and $\mu_H < 1$ *and party A has a type-*0 *culture while party B has a type-*1 *culture. Then party A*'s winning probability $p^A(\theta)$ $\left(= 1 - p^B(\theta)\right)$ *is given by* :

$$p^A(\theta) = \Pi\left([1 - 2\theta]\left[e^*(1) - \alpha e^*(1-\alpha) + \lambda(2\alpha - 1)^2 - \lambda\right]\right).$$

Formally, suppose that $\theta = 0$. If both parties were to decentralize, then $x = \alpha$ for both. However, $\pi = e^*(1)$ for A and $\pi = 0$ for B. Thus, B must centralize to compete with A. Then, voters get $\lambda + \alpha e^*(1-\alpha) + \chi$ under party B-rule and $\lambda(2\alpha - 1)^2 + e^*(1)$ under party A-rule. Therefore A has an electoral advantage (disadvantage) due its culture when $\theta = 0$ $(\theta = 1)$ and α is high.

Intuitively, this advantage comes from two sources: the ability to motivate party workers and better alignment with local interests. When $\theta = 0$, the party is decentralized and can take advantage of the motivated party workers; moreover the center and local party managers are aligned. Since the same θ shock hits both parties, party B has to centralize to compete, but this throws away local information and stops local party managers tailoring their campaigns to local interests. It also means that fewer party workers are motivated since the center is pushing against local party managers. While B has an advantage over A in that it compels greater coordination among party workers, such an advantage diminishes when α is close to 1.

On this view, the party culture most suitable for winning elections is context specific. Electoral success depends on the realization of θ in the short run and on β in the long run. Differences in political advantage due to party culture will be larger with stronger political competition, represented by a

higher density for popularity shock χ around its (zero) mean. This implies that any positive difference in $W^{A*} - W^{B*}$ maps into a larger difference in party A's probability of winning the election.[12]

Our earlier analysis can also explain why party cultures may not adapt to changed political circumstances, like permanent shifts in β which favor one party. Even though one party may want to modify its culture, this may be difficult, for the reasons explored above, giving it a lasting electoral disadvantage.

7 Final Remarks

We propose a model, where social identities held by overlapping generations of managers give rise to cultural dynamics. Our framework generates a range of insights into the interplay between organizational culture and design, with implications for performance. The model identifies the precise conditions for different long-run organizational cultures to emerge. Whether the organization is centralized or decentralized is endogenous and depends on internal conflicts of interest, which reflect tensions between its prevailing culture and its state-dependent leader objectives. We also apply these general ideas to four specific organizations.

Our framework could be extended in various ways. Hirschman (1970) famously emphasized three sources of organizational dynamics: exit, voice and loyalty. Here, we have focused on loyalty, as transmitted by social identity. But the model could also include exit and voice. Exit would reflect organizations under stress hiring outside managers with different cultural convictions to by-pass inside managers socialized into dysfunctional modes of behavior. It would be interesting to consider such hiring by embedding organizations in a market for managers. Voice would reflect managers having a more direct say in a centralized organization. For example, allowing senior managers to vote over the organization's mission – like the ρ chosen under centralization – would give an advantage to a majority culture. But one could study a variety of different governance mechanisms, including the selection of leaders and insiders' say in the selection process.

A wider set of issues about organizational governance and leadership could

[12] To see this concretely suppose that χ is uniform on $[-1/M, 1/M]$ then $\Pi(Z) = \frac{1}{2} + MZ$, assuming an interior solution. A higher density (more intense competition) then corresponds to a higher value of M.

be explored in our framework. For example, tasking a leader with a particular objective could have a long-run transformational effect. But a tasked leader may also create short-run unhappiness, by demotivating existing managers, as she attempts to transform the prevailing culture. Leader monitoring and evaluation will then be important – e.g., whether poor short-term performance is tolerated and not interpreted as the result of incompetence. Stories abound about leaders who attempt to change the culture of an organization, but are being edged out due to protests by disgruntled insiders or complaints by short-run oriented owners.

A richer theory of leaders would also be interesting. We have confined them to changing the authority structure. However, as already Weber (1922) emphasized, charismatic leaders can catalyze cultural change, quite apart from the sticks and carrots at their disposal. In terms of our model, this would somehow allow the leader to have a more direct effect on μ_t.

Finally, we have focused on how organizations adapt their design to endogenously changing values. We believe the idea of linking cultural and institutional change is a promising way of exploring societal dynamics in many contexts. In Besley and Persson (2017), we study how evolving democratic values interact with reforms of democratic institutions. More research on the interplay between formal rules and cultural values remains scarce could help us better understand the drivers of economic success and failure.

References

[1] Acemoglu, Daron [2003], "Why not a Political Coase Theorem? Social Conflict, Commitment, and Politics," *Journal of Comparative Economics* 31, 620-652.

[2] Aghion Philippe and Jean Tirole [1997], "Formal and Real Authority in Organizations," *Journal of Political Economy* 105, 1-29.

[3] Ahmad, Junaid, Shantayanan Devarajan, Stuti Khemani, and Shekhar Shah [2005], "Decentralization and Service Delivery," Policy Research Working Paper 3603, The World Bank: Washington, D.C.

[4] Akerlof, George [1976], "The Economics of Caste, the Rat-Race and Other Woeful Tales," *Quarterly Journal of Economics* 90, 599-617.

[5] Akerlof, George, and Rachel Kranton [2000], "Economics and Identity," *Quarterly Journal of Economics* 115, 715–753

[6] Akerlof, George and Rachel Kranton [2010], *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, Princeton, NJ: Princeton University Press.

[7] Alesina, Alberto and Paola Giuliano [2015], "Culture and Institutions", *Journal of Economic Literature* 53, 898-944.

[8] Alger, Ingela and Jörgen Weibull [2013], "Homo Moralis – Preference Evolution under Incomplete Information and Assortative Matching," *Econometrica* 81, 2269–2302.

[9] Alonso, Ricardo, Wouter Dessein, and Niko Matouschek [2008], "When Does Coordination Require Centralization?", *American Economic Review* 98, 145-179.

[10] Ashforth, Blake E. and Fred Mael [1989], "Social Identity Theory and the Organization," *The Academy of Management Review* 14, 20-39.

[11] Bandiera, Oriana, Stephen Hansen, Raffaella Sadun, and Andrea Prat [2016], "CEO Behavior and Firm Performance," mimeo LSE.

[12] Benabou, Roland [2013], "Groupthink: Collective Delusions in Organizations and Markets," *Review of Economic Studies* 80, 429-462.

[13] Besley, Timothy and Maitreesh Ghatak [2005], "Competition and Incentives with Motivated Agents," *American Economic Review* 95, 616-636.

[14] Besley, Timothy and Torsten Persson [2017], "Democratic Values and Institutions," mimeo, LSE and IIES.

[15] Bisin, Alberto and Thierry Verdier [2011], "The Economics of Cultural Transmission and Socialization", Chapter 9 in Benhabib, Jess (ed.), *Handbook of Social Economics*, Volume 1A, Elsevier.

[16] Bloom, Nicholas, Renata Lemos, Raffaella Sadun, and John Van Reenen [2015], "Does Management Matter in Schools?", *Economic Journal* 125, 647–674.

[17] Bloom, Nicholas, Raffaella Sadun, and John Van Reenen [2012], "The Organization of Firms Across Countries," *Quarterly Journal of Economics* 127, 1663-1705.

[18] Bloom, Nicholas, Raffaella Sadun, and John Van Reenen [2014], "Does Management Matter in Healthcare," unpublished typescript.

[19] Bolton, Patrick and Joseph Farrell [1990], "Decentralization, Duplication, and Delay," *Journal of Political Economy* 98, 803-826.

[20] Boyd, Robert and Peter Richerson [1985], *Culture and the Evolutionary Process*, Chicago, IL: University of Chicago Press.

[21] Cavalli-Sforza, Luigi L. and Marcus Feldman [1981], *Cultural Transmission and Evolution*, Princeton, NJ: Princeton University Press.

[22] CHKS [2012], "What Makes a Top Hospital? Report 4: Organizational Culture," available at http://www.chks.co.uk/userfiles/files/CHKS_2012_WMATH_4%20-%20final.pdf.

[23] Coase, Ronald H. [1937], "The Nature of the Firm," *Economica* 4, 386-405.

[24] Cox, Gary W. and Matthew McCubbins [2007], *Legislative Leviathan: Party Government in the House*, 2nd edition, Cambridge University Press.

[25] Cyert, Richard and James G. March [1963], *A Behavioral Theory of the Firm,* Oxford: Wiley-Blackwell.

[26] Dessein, Wouter and Andrea Prat [2017], "Organizational Capital, Corporate Leadership, and Firm Dynamics," Columbia Business School Research Paper No. 17-89.

[27] Dessi, Roberta [2008], "Collective Memory, Cultural Transmission and Investments", *American Economic Review* 98, 534–560.

[28] Gioia, Dennis A. and James B. Thomas [1996], "Identity, Image, and Issue Interpretation: Sensemaking during Strategic Change in Academia," *Administrative Science Quarterly* 41, 370-403.

[29] Gelfand, Michele J., Lisa M. Leslie, and Kirsten M. Keller [2008], "On the Etiology of Conflict Cultures," *Research in Organizational Behavior* 28, 137–166.

[30] Grief, Avner [1994], "Cultural Beliefs and the Organization of Society: A Theoretical and Historical Reflection on Collectivist and Individualist Societies", *Journal of Political Economy* 102, 912-950.

[31] Grossman, S. and Oliver D. Hart [1986], "The Costs and Benefits of Ownership: A Thory of Vertical and Lateral Integration," *Journal of Political Economy* 94, 691-719.

[32] Guiso, Luigi, Paola Sapienza, and Luigi Zingales [2006], "Does Culture Affect Economic Outcomes?", *Journal of Economic Perspectives* 20, 23–48.

[33] Guiso, Luigi, Paola Sapienza, and Luigi Zingales [2013], "The Values of Corporate Culture," *Journal of Financial Economics* 117, 20-36.

[34] Güth,Werner [1995], "An Evolutionary Approach to Explaining Cooperative Behavior by Reciprocal Incentives," *International Journal of Game Theory*, 323–344.

[35] Güth, Werner and Menahem E. Yaari [1992], "Explaining Reciprocal Behavior in Simple Strategic Games: An Evolutionary Approach," in Witt, Ulrich (ed.) *Explaining Process and Change: Approaches to Evolutionary Economics*, Ann Arbor, MI: University of Michigan Press.

[36] Hart, Oliver D. and Bengt Holmström [2010], "A Theory of Firm Scope," *Quarterly Journal of Economics* 125, 483-513.

[37] Hermalin, Benjamin [2001], "Economics and Corporate Culture," in S. Cartwright et al (eds.), *The International Handbook of Organizational Culture and Climate,* Chichester England: John Wiley & Sons, Ltd.

[38] Hirschman, Albert O. [1970], *Exit, Voice and Loyalty*, Cambridge, MA: Harvard University Press.

[39] Hodgson, Geoffrey M. [1996], "Corporate Culture and the Nature of the Firm," in John Groenewegen, (ed.), *Transaction Cost Economics and Beyond*, Boston, MA: Kluwer Academic Press.

[40] Hofstede, Geert [1984], *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations,* New York, NY: Sage Publications.

[41] Hofstede, Geert, Gert Jan Hofstede, and Michael Minkov [2010], *Cultures and Organizations: Software of the Mind*, New York, NY: Mc Graw Hill, Third Edition.

[42] Inglehart, Ronald, Miguel Basanez, Jaime Diez-Medrano, Loek Halman, and Ruud Luijkx (eds.) [2004], *Human Beliefs and Values: A Cross-cultural Sourcebook Based on the 1999–2002 Values Surveys*, Mexico City: Siglo XXI.

[43] Kreps, David M. [1990], "Corporate Culture and Economic Theory," in Alt, James and Kenneth A. Shepsle (eds.), *Perspectives on Positive Political Economy*, Cambridge, England: Cambridge University Press.

[44] Lazear, Edward P. [1995], "Corporate Culture and the Diffusion of Values," in Horst Siebert (ed.), *Trends in Business, Organization*, Tübingen, Germany: J.C.B.

[45] Lucas, Robert E. Jr. [1978], "On the Size Distribution of Business Firms", *Bell Journal of Economics* 9, 508-523.

[46] Mills, D. Quinn [1996], "The Decline and Rise of IBM," *Sloan Management Review* 37, 78-82.

[47] Peters, Tom and Robert H. Waterman Jr. [1982], *In Search of Excellence: Lessons from America's Best-Run Companies*, London, UK: Harper Collins.

[48] Prendergast, Canice [2007], "The Motivation and Bias of Bureaucrats," *American Economic Review* 97, 180-196.

[49] Rogoff, Kenneth [1985], "The Optimal Degree of Commitment to an Intermediate Monetary Target," *Quarterly Journal of Economics* 100, 1169-1189.

[50] Sandholm, William H. [2010], *Population Games and Evolutionary Dynamics*, MIT Press, Cambridge, MA.

[51] Schein, Edgar H. [1983], "The Role of the Founder in Creating Organizational Culture," *Organizational Dynamics* 12, 13–28.

[52] Schein, Edgar H. [1990], "Organizational Culture," *American Psychologist* 45, 109–119.

[53] Schein, Edgar H. [1996], "Three Cultures of Management: The Key to Organizational Learning," *Sloan Management Review* 38, 9–20.

[54] Shleifer, Andrei and Lawrence Summers [1988], "Breach of Trust in Hostile Takeovers," in Auerbach. Alan (ed.), *Corporate Takeovers: Causes and Consequences*, Chicago, IL: University of Chicago Press.

[55] Tirole, Jean [1994], "The Internal Organization of Government," *Oxford Economic Papers* 46, 1-29.

[56] Tirole, Jean [1996], "A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality)," *Review of Economic Studies* 63, 1–22.

[57] Vickers, John [1985], "Delegation and the Theory of the Firm," *Economic Journal* 95, Supplement: Conference Papers, 138-147.

[58] Weber, Max [1922], "The Nature of Charismatic Authority and its Routinization', in *Theory of Social and Economic Organization*, 1947 translation by Anderson, A. R. and Talcott Parsons.

[59] Whyte, William H. [1956], *The Organization Man*, New York, NY: Simon & Schuster.

[60] Williamson, Oliver E. [1979] "Transaction-Cost Economics: The Governance of Contractual Relations," *Journal of Law and Economics* 22, 233-261.

[61] Willis, Eliza, Christopher da C. B. Garman, and Stephan Haggard [1999], "The Politics of Decentralization in Latin America," *Latin American Research Review* 34, 7-56.

[62] Wilson, James Q. [1989], *Bureacracy: What Government Agencies Do and Why They Do It*, New York, NY: Basic Books.

[63] Wolcott, Robert C. and Michael J. Lippitz [2007], "The Four Models of Corporate Entrepreneurship," *Sloan Management Review* 49, 75-82.

Appendix

A Derivation of $q(\mu)$ and $Q(\Delta)$

Culture in the sense of the share of type-0 managers evolves over time. We have deliberately simplified by assuming that all upper-tier managers leave each period, and all lower-tier managers are promoted. Therefore, μ_{t+1} is pinned down by the way types are transmitted from upper-tier to lower-tier managers in period t.

A microfounded socialization mechanism One possible transmission mechanism builds on direct and indirect socialization. Let us assume that being randomly matched with an upper-tier manager at stage 4 of the period involves a mentoring component. This mentoring helps determine the lower-tier manager's type, which becomes relevant once he is promoted.

If a lower-tier manager is mentored by a type-0 manager, which happens with probability μ_t, we assume that he may acquire the same type, depending on the relative expected fitness of holding the two types as a senior manager in the next period. Specifically, let $\Delta\left(\mu_{t+1}\right) = E[v(0,\omega,\theta) - v(1,\omega,\theta)]$ be tomorrow's expected-utility difference between having type 0 and type 1 with a share of μ_{t+1} type-0 managers in the organization. Then, a lower-tier manager becomes type 0 through mentoring if:

$$\Delta\left(\mu_{t+1}\right) + \eta \geq 0,$$

where η is a mean-zero, symmetrically distributed idiosyncratic shock with continuous distribution function $G(\cdot)$. Thus the probability that that a new recruit mentored by a type-0 upper-tier manager himself becomes type 0 is just $G\left(\Delta\left(\mu_{t+1}\right)\right)$.

If such direct socialization fails, the lower-tier manager may still be indirectly socialized by observing and learning from other managers. The probability of indirectly becoming type 0 depends monotonically on the average fraction of such types in the organization, a kind of social learning postulated in much of the cultural-evolution literature. Assuming a linear relation, the probability of indirect socialization becomes $\left(1 - G\left(\Delta\left(\mu_{t+1}\right)\right)\right)\mu_t$.

Adding these expressions, the overall probability that a new recruit who is matched with a type-0 upper-tier manager himself acquires this type is:

$$G\left(\Delta\left(\mu_{t+1}\right)\right) + \left(1 - G\left(\Delta\left(\mu_{t+1}\right)\right)\right)\mu_t. \tag{18}$$

If a new lower-tier manager is matched with and mentored by a type-1 upper-tier manager, which happens with probability $1-\mu_t$, he is never directly socialized into becoming type 0. On the other hand he is socialized into being type 1 if

$$\Delta\left(\mu_{t+1}\right)+\eta\leq 0.$$

Thus, $\left(1-G\left(\Delta\left(\mu_{t+1}\right)\right)\right)$ is the proportion of type-1 managers coming from such matches. The fraction $G(\Delta\left(\mu_{t+1}\right))$ of lower-tier managers who do not become type 1 in this way, can – as above – indirectly become type 0 depending on the aggregate fraction of type-0 upper-tier managers in the organization. The resulting probability of becoming a type-0 manager is $G\left(\Delta\left(\mu_{t+1}\right)\right)\mu_t$.

Multiplying (18) with μ_t, $G\left(\Delta\left(\mu_{t+1}\right)\right)\mu_t$ with $1-\mu_t$, and adding the resulting expressions, we can write the equation of motion for the share of type-0 managers as

$$\begin{aligned}\mu_{t+1} &= \mu_t\left[G\left(\Delta\left(\mu_{t+1}\right)\right)+\left(1-G\left(\Delta\left(\mu_{t+1}\right)\right)\right)\mu_t\right]+\left(1-\mu_t\right)G\left(\Delta\left(\mu_{t+1}\right)\right)\mu_t\\ &= \mu_t+\left(1-\mu_t\right)\mu_t 2\left[G\left(\Delta\left(\mu_{t+1}\right)\right)-\frac{1}{2}\right]. \qquad (19)\end{aligned}$$

The expression on the right-hand side in consistent with the assumptions made about $Q(\Delta)$ and $q(\mu_t)$ made in the text. Note that here Δ depends on tomorrow's culture μ_{t+1}.

A Replicator Dynamic An even simpler approach to the replicator dynamic is to suppose that there is matching between a young and old manager. The probability a type 1 senior manager does not convert the young manager to being a type 1 is

$$\begin{aligned}\rho_{10}\left(\mu_t\right) &= \frac{E(v(0,\omega,\theta):\mu_t)-\left[\mu_t E(v(0,\omega,\theta):\mu_t)+\left(1-\mu_t\right)E(v(1,\omega,\theta:\mu_t))\right]}{2\bar{\Delta}}\\ &= -\left(1-\mu_t\right)\left[\frac{\Delta\left(\mu_t\right)}{2\bar{\Delta}}\right]\end{aligned}$$

and the probability that a type 0 manager fails to convert the lower tier manager with whom he is matched to be type 0 is

$$\begin{aligned}\rho_{01}\left(\mu\right) &= \frac{E(v(1,\omega,\theta):\mu_t)-\left[\mu_t E(v(0,\omega,\theta):\mu_t)+\left(1-\mu_t\right)E(v(1,\omega,\theta:\mu_t))\right]}{2\bar{\Delta}}\\ &= \mu_t\left[\frac{\Delta\left(\mu_t\right)}{2\bar{\Delta}}\right]\end{aligned}$$

with $\Delta(0) > -\bar{\Delta}$ and $\Delta(1) < \bar{\Delta}$. Then

$$\begin{aligned}\mu_{t+1} &= \mu_t(1-\rho_{10}(\mu_t)) + \rho_{01}(\mu_t)(1-\mu_t) \\ &= \mu_t + \mu_t(1-\mu_t)\frac{\Delta(\mu_t)}{\bar{\Delta}}.\end{aligned}$$

To see the equivalence to the replicator dynamic note that

$$\begin{aligned}\mu_{t+1} &= \mu_t + \mu_t(1-\mu_t)\frac{\Delta(\mu_t)}{\bar{\Delta}} \\ &= \mu_t + \mu_t\frac{\{E(v(0,\omega,\theta):\mu_t) - [\mu_t E(v(0,\omega,\theta):\mu_t) + (1-\mu_t)E(v(1,\omega,\theta:\mu_t))]\}}{\bar{\Delta}}.\end{aligned}$$

Sandholm (2010) discusses a range of micro-foundations for the replicator dynamic in population games.

A.1 Proofs of Lemmas and Propositions

Proof of Lemma 1 Given that the first and third arguments are the same in $\Pi(.,.,.)$, only the second argument matters. So $\rho(\theta)$ depends on maximizing average profits. Note that with centralization and $\theta = 0$, we have $\rho(0) = 0$ if

$$\alpha\pi(0,0) + (1-\alpha)\pi(1,0) \geq \alpha\pi(1,0) + (1-\alpha)\pi(0,0).$$

If $\theta = 1$, then we have $\rho(1) = 0$ if

$$\alpha\pi(1,1) + (1-\alpha)\pi(0,1) \geq \alpha\pi(0,1) + (1-\alpha)\pi(1,1).$$

Both inequalities hold strictly, since $\alpha \geq \frac{1}{2}$, $\pi(0,0) > \pi(1,0)$ and $\pi(1,1) > \pi(0,1)$.

Proof of Proposition 1 Let $\theta = 0$ and define

$$\begin{aligned}\Pi\left(\lambda(2[\mu_H\alpha + (1-\mu_H)(1-\alpha)] - 1)^2, \mu_H\pi(0,0) + (1-\mu_H)\pi(1,0), e\right) \\ = \Pi(\lambda, \alpha\pi(0,0) + (1-\alpha)\pi(1,0), e),\end{aligned}$$

which must have $\mu_H \geq \alpha \geq 1/2$. Because the LHS is increasing in μ, part 1 follows.

Let $\theta = 1$ and define

$$\Pi\left(\lambda(2[\mu_L\alpha + (1-\mu_L)(1-\alpha)] - 1)^2, (1-\mu_L)\,\pi\,(1,1) + \mu_L\pi\,(0,1)\,, e\right) =$$
$$= \Pi\left(\lambda, \alpha\pi\,(1,1) + (1-\alpha)\,\pi\,(0,1)\,, e\right),$$

which must have $1 - \mu_L \geq \alpha \geq 1/2$. Because the LHS is decreasing in μ, part 2 follows.

Proof of Lemma 2 To prove this, we start from

$$\mu_{t+1} - \mu_t = q(\mu_t)Q(\Delta_t).$$

If $\Delta(\mu)$ is globally increasing, $q(\mu_t) > 0$, and $\Delta(\hat{\mu}) = 0$, we must have $\mu_{t+1} - \mu_t \geq 0$ for all $1 \geq \mu \geq \hat{\mu}$, while $\mu_{t+1} - \mu_t < 0$ for all $0 \leq \mu < \hat{\mu}$. The interior steady state is thus unstable. Moreover, $\Delta(\mu)$ globally increasing implies $Q(\Delta(1)) \geq 0 \geq Q(\Delta(0))$. This implies that the steady states at $\mu = 0$ and $\mu = 1$ are stable.

Proof of Lemma 3 From the definitions in the text, we can guarantee that $\Delta(\mu)$ is globally increasing if (i) $\hat{\delta}(\mu_H) \geq \delta_H(\mu_H)$ (ii) $\hat{\delta}(\mu_L) \leq \delta_L(\mu_L)$, and (iii) $\delta_L(\mu)$ increasing for $\mu \geq \alpha$. Define

$$\Omega_H(\mu) = [\xi\,[2\mu - 1]]\,e^*(1) - e^*(\nu(\mu))\,[2\alpha - 1 + \xi(\mu + \alpha - 1)]$$

and note that (i) is equivalent to $\Omega_H(\mu_H) \geq 0$. This condition will hold for

$$\xi \geq \frac{e^*(\nu(\mu))\,[2\alpha - 1]}{[(2\mu - 1)e^*(1) - e^*(\nu(\mu))\,(\mu + \alpha - 1)]}.$$

Next, define

$$\Omega_L(\mu) = e^*(1 - \nu(\mu))[1 - 2\alpha + (\mu - \alpha)\xi] - [\xi\,[2\mu - 1]]\,e^*(1)$$

and note that (ii) is equivalent to $\Omega_L(\mu_L) > 0$. This condition holds if

$$\xi \geq \frac{e^*(1 - \nu(\mu_L))[2\alpha - 1]}{[1 - 2\mu_L]\,e^*(1) - e^*(1 - \nu(\mu_L))(\alpha - \mu_L)]}.$$

So we need ξ to satisfy:

$$\xi \geq \max\left\{\frac{e^*(1 - \nu(\mu_L))[2\alpha - 1]}{[1 - 2\mu_L]\,e^*(1) - e^*(1 - \nu(\mu_L))(\alpha - \mu_L)]}, \frac{e^*(\nu(\mu_H))\,[2\alpha - 1]}{e^*(1)\,[2\mu_H - 1] - e^*(\nu(\mu_H))\,(\mu_H + \alpha - 1)]}\right\} \tag{20}$$

Finally, we would like $\delta_L(\mu)$ to be increasing for all $\mu \geq \mu_H$. This is the case if

$$\begin{aligned} & e^*(1-\nu(\mu))\xi] + (1-2\alpha)\frac{\partial e^*(1-\nu(\mu))}{\partial \nu}[1-2\alpha+(\mu-\alpha)\xi] \\ = \; & e^*(1-\nu(\mu))\xi] + (1-2\alpha)^2\frac{\partial e^*(1-\nu(\mu))}{\partial \nu}[1+\frac{(\mu-\alpha)}{1-2\alpha}\xi] > 0. \end{aligned}$$

For this condition to hold at large enough ξ, we need that

$$e^*(1-\nu(\mu)) + (1-2\alpha)(\mu-\alpha)\frac{\partial e^*(1-\nu(\mu))}{\partial \nu} > 0.$$

This condition is Assumption 1.

Proof of Proposition 2 In Case 3, β is such that the leader fluctuates in their views often enough for there to be multiple stable steady states. Let

$$\varphi(\mu,\beta) = \beta e^*(\nu(\mu))[2\alpha-1+\xi(\mu+\alpha-1)]+(1-\beta)e^*(1-\nu(\mu))[1-2\alpha+(\mu-\alpha)\xi].$$

Note that $\varphi(\mu,\beta)$ is increasing in μ and $\varphi(\tilde{\mu}(\beta),\beta) = 0$. Under Lemma 2, $\Delta(\mu)$ is increasing in μ. Suppose there exists β such that $\tilde{\mu}(\beta) \in [\mu_L, \mu_H]$. Then if $\mu > \tilde{\mu}(\beta)$ we have $\Delta(\mu) > 0$, and if $\mu < \tilde{\mu}(\beta)$ we have $\Delta(\mu) < 0$.

Proof of Proposition 3 In general, with $\mu = 1$ the long-run expected payoff is

$$\beta\Pi(\lambda, \pi(0,0), e^*(1)) + (1-\beta)\Pi(\lambda, [\alpha\pi(1,1) + (1-\alpha)\pi(0,1)], e^*(1-\alpha)).$$

With $\mu = 0$ it is instead

$$\beta\Pi(\lambda, [\alpha\pi(0,0) + (1-\alpha)\pi(1,0)], e^*(1-\alpha)) + (1-\beta)\Pi(\lambda, \pi(1,1), e^*(1)).$$

The payoff is higher (lower) with $\mu = 1$ ($\mu = 0$) if and only if

$$\begin{aligned} & \beta[\Pi(\lambda, \pi(0,0), e^*(1)) - \Pi(\lambda, [\alpha\pi(0,0) + (1-\alpha)\pi(1,0)], e^*(1-\alpha))] > (<) \\ & (1-\beta)[\Pi(\lambda, \pi(1,1), e^*(1)) - \Pi(\lambda, [\alpha\pi(1,1) + (1-\alpha)\pi(0,1)], e^*(1-\alpha))]. \end{aligned} \tag{21}$$

In the multiplicative case, this boils down to

$$\begin{aligned} & \beta[\pi(0,0)e^*(1) - [\alpha\pi(0,0) + (1-\alpha)\pi(1,0)]e^*(1-\alpha)] > (<) \\ & (1-\beta)[\pi(1,1)e^*(1) - [\alpha\pi(1,1) + (1-\alpha)\pi(0,1)]e^*(1-\alpha)] \end{aligned}$$

which yields the condition in the proposition.

Proof of Proposition 4 See the text in Subsection 6.2.

Proof of Proposition 5 The result follows from observing that, with $\mu^A = 1$ and $\mu^B = 0$, $\theta = 0$ implies

$$W^{A*} - W^{B*} = e^*(1) - \alpha e^*(1-\alpha) + \lambda(2\alpha - 1)^2 - \lambda.$$

This follows as party A will decentralize and have $x_A = \alpha$, while party B will centralize and set $\rho(0) = 0$ with effort $e^*(1 - \nu(0)) = e^*(1-\alpha)$ and a fraction α of local parties aligned with the state. A parallel argument says that with $\theta = 1$, then

$$W^{A*} - W^{B*} = \alpha e^*(1-\alpha) - e^*(1) + \lambda - \lambda(2\alpha - 1)^2.$$

Putting these together yields the result.