Joint Dynamics of Organizational Culture, Design, and Performance*

Timothy Besley
LSE and CIFAR

Torsten Persson
IIES and CIFAR

June 2017

Abstract

This paper examines organizational cultures created through emerging values and how these cultures interact with a key aspect of organizational design, namely how far key decisions are delegated. We model culture as adopting specific social identities in an organization which affect the way that management decisions are taken. Using a simple dynamic model of socialization based on the relative payoffs of each type, the paper investigates the conditions under which different cultures become dominant. Our general model delivers insights into the emergence of dysfunctional cultures and resistance to change. We apply this model to the behavior of bureaucracies, firms, and political parties.

*We are grateful to Philippe Aghion, Maitreesh Ghatak, Bob Gibbons, Oliver Hart and Jay Lee as well as participants in meetings with the IOG program of the Canadian Institute for Advanced Research, the ERINN network, and the Growth and Institutions program at Tsinghua University for helpful discussions and comments.

1 Introduction

Research on public as well as private organizations has focused on understanding the drivers of their performance. Despite a widespread acknowledgement that organizational culture may be critical among these drivers, there is little agreement on how to capture this concept in economic models. A common approach is to model culture as shaping the beliefs that govern 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 organization, 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.

In this paper, we model culture as internalized values that influence behavior. This idea accords with the well-known approach 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 is transmitted to incoming members of the organization through a process of socialization, which creates group identities among the organization's managers. These identities, in turn, affect the organization's choice of mission and hence its performance. By stressing social identity, we follow Ashforth and Mael (1989) in the organizational-behavior literature and Akerlof and Kranton (2000) in the economics literature. Social-identity theory – which was developed among sociologists – assumes that an individual's identity is derived 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 other group members and become more likely to 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. First, for a given set of identities in the organization, its leader chooses how much discretion to pass down to the next layer of management. When the culture embodied

in these identities is congruent with the leader's objective, she is more likely to decentralize as in Aghion and Tirole (1997). Second, decentralization decisions affect the dynamics of organizational culture. The latter follow an evolutionary process of socialization, where the dominance of a culture depends on its relative organizational fitness, as in Boyd and Richerson (1985). Because of these features, 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 organization design; decentralization has a natural upside in taking advantage of local information, but this requires alignment of the prevailing culture with the organization's objective. Second, we show that 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 cultures become entrenched. Third – and partly as a corollary of the second phenomenon – dysfunctional cultures can emerge in the long run. That is, an organization's culture can become entrenched, even though it does not serve the organization's purpose according to standard performance criteria. Fourth, cultures form basins of attraction, which make organizations less responsive to shocks to the environment in which they operate.

To breathe life into our general model, we apply it to four specific kinds of organizations. One application concerns public bureaucracies, where we stress dilemmas of top-down control and differential performance by different units. Another one is about private firms, and demonstrates how observed correlations between productivity, culture (management style), and organizational form can arise endogenously over time. We also ask if stronger market conditions weed out dysfunctional cultures. Our third application illustrates how a strong organizational culture may become a barrier to innovation when a firm tries to adapt to new market conditions. The 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 a motivating example. In Section 4, we develop 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

The economics literature on corporate culture is too vast to survey here. We refer the reader to the excellent survey by Hermalin (2001), which identifies various strands of the literature. One important approach, taken by Kreps (1990) and others, is to regard culture as a belief-based norm in a game played by overlapping generations of agents, where cooperation is sustained against the threat of poor future performance. A different approach, taken by Hodgson (1996) and Lazear (1995), is more similar to our own in stressing how different types evolve within an organization.

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 which have beliefs more supportive of cooperation. An alternative approach taken by Akerlof (1976) and Akerlof and Kranton (2000) see manifestations of culture in preferences that drive individual behavior. Our modeling builds on the latter approach.

We also build on models of cultural evolution, inspired by research beginning with Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985). Studies of socialization and cultural economics has grown in recent years, and 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).

Empirical measurement of cultural differences have flourished, but largely outside of economics. For example, Hofstede (1984) began a body of research on international comparisons of organizational cultures.[1] The well-known World Values Survey was developed as a means of examining cultural differences (see Inglehart et al, 2004). Nowadays, however, empirical studies of culture have also become extensive in economics (see, Alesina et al, 2015 and Guiso et al, 2006 for overviews). While these ideas have mostly been applied to individuals, they have also been applied to firms. For example, Guiso et al (2015) argue that corporate cultures that include integrity are likely to improve performance.

A large literature in business economics and sociology studies conflicts of interest inside firms, with many authors taking Cyert and March (1963) as

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

a starting point. Economists modeling such ideas have asked how conflicting interests shape 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 the informational benefits of delegation are weighed against the value of coordination. The resulting literature has influenced empirical studies of firm behavior. In the same tradition, Bloom et al (2012) look empirically at decentralization by firms across countries, finding productivity gains from decentralization associated with greater levels of trust. Bandiera et al (2016) examine how CEOs use their scarce time, and find the largest differences regarding the direct involvement in production vs. coordination with high-level executives. In our setting, conflicts of interest and delegation arise endogenously over time, as a result of the interaction between the 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 the statement by Wolcott and Lippitz (2007) who suggest that

> "Unless a company if 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). That line of work is more influenced by sociology, psychology, and anthropology than by economic approaches. Researchers in this field have debated at length how organizational cultures are created, where 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 conflicts that can emerge once cultures have become established (see, e.g., Gelfand et al, 2015). Our focus on the role of group identities provides a bridge from these ideas to a more economic perspective.

3 Innovation and Culture at IBM

To frame the ideas to follow, we discuss a concrete example with several interesting features. It concerns the case of IBM which has been the subject of

many studies of organizational culture, including the classic work by Hofstede (1984). The company's appearance in key case studies of organizational dynamics and culture partly reflects the strong ethos and charisma of CEO Thomas J. Watson. Leading textbooks on the origins of corporate success, like Peters and Waterman (1982), have also featured IBM as a prominent example and argued

> "(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 reinforces the importance of values and the need to encourage and motivate employees to focus on things which have the highest return to the organization.

In the 1950s, IBM became 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 missing the fast-growing mini-computer market during the 1970s and losing out to its rivals. In 1979, this led *Business Week* to suggest that IBM was a "stodgy, mature company", a view supported by a decline in IBM's stock price by around 20%. In an effort to avoid falling behind in the new personal-computer industry, the firm began working on the now-famed IBM PC, prompting the well-known 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 faced in shifting focus away from mainframes to networks and personal computing. Mills (1996) discusses this experience based on interviews with IBM management. He explicitly emphasizes the need to balance 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 blames the role of culture in limiting IBM's capacity to respond:

> "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 has three features which are key to our model. First, the organizational culture is important in engendering a sense of belonging and a common interest among groups of employees. Second, once it is entrenched, a culture can limit an organization's adaptability in the wake of changing priorities and market conditions. Third, when top leaders clash with a prevailing management culture, this poses the question how far to centralize decision making.

While the IBM example is only illustrative, we believe the same features arise in many contexts. For example, similar concerns are now being expressed about the prospects for Google trying to adapt to greater competition and new product lines, such as taking on Facebook and adapting to the use of mobile apps.

4 Basic Framework

Overview We study an organization with a three-tier hierarchy: a leader, a set of senior managers and a set of junior managers. Tomorrow's senior managers are drawn from today's junior managers. The leader wants to run the organization to suit prevailing conditions, which affect how well the organization does. However, if she decentralizes, the outcome depends on senior managers' mission-driven preferences that embody a certain culture or ethos. While it may be tempting to centralize key decisions, this could de-motivate managers.

Organizational culture takes the form of preferences for a particular mission. In the IBM example, this could represent projects that benefit the

development and sales of mainframe systems rather than PCs. Such preferences are transmitted from one generation to the next via a socialization process. In this process, the likelihood of acquiring a particular "tribal" preference depends (probabilistically) on current perceptions about relative payoffs. Group membership – organizational culture – thus evolves with these payoffs.

When leaders choose whether or not to centralize, they know the tribal composition of its management, but not precisely which manager belongs to which group. While centralization guarantees adoption of a particular mission, it throws away information about local conditions. The centralization decision also affects cultural dynamics: whether senior managers have discretion affects the relative payoff of tribes. These two-way feedbacks create a coevolutionary process between organizational form and culture.

Organizations and states of the world An organization has a continuum of divisions. These divisions have unit measure and are indexed by $\omega \in [0,1]$. Each division has to make a design choice, denoted by $\rho(\omega,\theta) \in \{0,1\}$ where $\theta \in \{0,1\}$ is the aggregate state. This aggregate state and the relationship between $\rho(\omega,\theta)$ and a state for local production conditions $\sigma(\omega,\theta) \in \{0,1\}$ jointly determine firm performance (as further discussed below). In the IBM example, we could think about state $\theta = 0$ ($\theta = 1$) as condition favoring mainframes (PCs). The probability of $\theta = 0$ is β. Hence, the value of β captures the predictability of the organization's environment. When β is close to either 0 or 1 it is highly predictable, whereas a β closer to $\frac{1}{2}$ entails a more unpredictable environment.

For both values of θ, $\alpha \geq \frac{1}{2}$ is the probability that any division has $\sigma(\omega,\theta) = \theta$. Hence, $\alpha \in \left[\frac{1}{2},1\right]$ gauges how well technology, demand, or cost is aligned across divisions.

Members of the organization The organization has a leader. She represents the organization's ultimate principal(s) – the owners of the firm, the ministry, or customers of the bureaucracy, or the voters of the party – and shares their preferences. The leader observes θ but not $\sigma(\omega,\theta)$. Further, she chooses the organizational form $o \in \{d,c\}$, where d stands for decentralized and c for centralized. In a centralized organization, she also chooses a unique local action $\rho(\theta) \in \{0,1\}$ which is binding for all divisions ω.

Each division is staffed by an upper-tier manager, indexed by U, and a

lower-tier manager, indexed by L. These sets of managers have two-period overlapping lives, such that this period's upper-tier managers are replaced in the next period by this period's lower-tier managers.

Each upper-tier manager observes local conditions $\sigma(\omega, \theta)$ in his own division (as well as θ). The upper-tier manager thus has better local information than the leader, but the latter's information disadvantage diminishes in alignment parameter α. In a decentralized organization, the leader delegates design choice ρ to her upper-tier managers.

Leader payoffs The leader maximizes the payoff

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

The payoff function $\Pi(\cdot)$ is increasing in each term. The first term $(2x-1)^2$ is a measure of coordination in the organization, where x is the (maximum) share of divisions that takes the same action ρ.[2] This term is maximized (at 1) when every division makes the same choice (either $\rho = 1$ or $\rho = 0$). Parameter λ indexes the importance of coordination gains. Thus, greater coordination is always valuable, *ceteris paribus*. This way of capturing the benefits of coordination is similar to that in the literatures on the scope of the firm (Hart and Holmström, 2010) and 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 divisional decisions to local conditions. Here, $\pi|\rho(\omega,\theta) - \sigma(\omega,\theta)|, \theta)$ is the payoff to alignment in division ω.Throughout we assume that

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

This says that matching the local state with the local action is always optimal in state $\theta = 0$ and never optimal in state $\theta = 1$. In the IBM example, state 0 favoring mainframes makes organization payoffs the highest when the managers choose projects more directed to mainframes by setting $\rho(\omega,\theta) - \sigma(\omega,\theta) = 0$.

[2] The symmetry of the two states in the model means that we could equivalently define x as the fraction of divisions that set $\rho = 1$.With the partuicular functional form $(2x-1)^2$, this would give identical results to the "max" formulation.

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

A special case of the leader's payoff, which we use in some of the applications in Section 6, occurs when

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

which we refer to as "the multiplicative case".

The leader observes the aggregate state θ and the organization's culture μ (see below), but does not observe the local state $\sigma\left(\omega,\theta\right)$ and the type $\tau(\omega)$ (see below) in each division. She always choose the organizational form $o(\theta)$.

Upper-tier managers Upper-tier managers in division ω can be one of two possible types $\tau(\omega) \in \{0,1\}$. A share μ_t of divisions have type $\tau\left(\omega\right)=0$ and lexicographically prefer $\rho(\omega,\theta) = \sigma\left(\omega,\theta\right)$, while $1-\mu_t$ have a type $\tau\left(\omega\right)=1$ manager and prefer $\rho(\omega,\theta) = 1-\sigma\left(\omega,\theta\right)$. In the remainder of the paper, we refer to the distribution of types – captured by μ_t – as the *organizational culture*. Note that in *state* $\theta=0$ the leader's preferences are aligned with the preferences of *type* $\tau=0$ managers but clashes with those of type $\tau=1$ managers, and vice versa in state $\theta=1$. At a given point in time, these alignments or conflicts of interest – due to the juxtaposition of states and types – will determine the leader's willingness to decentralize project choices, so as to take advantage of local (divisional) information, precisely as in Aghion and Tirole (1997). In our model, the incentive to decentralize will not stay constant, however, but change over time if organizational culture μ_t is evolving. We study that evolution in the next section.

Identifying with a certain type gives upper-tier managers a form of "tribal" preferences: they care not only about their own payoff, but also about the average payoff of all other upper-tier managers who have the same type. Thus they internalize the payoffs of the whole group of managers who belong to their tribe.

The direct payoff to an upper-tier manager is given by

$u(\left|\rho(\omega,\theta-\sigma\left(\omega,\theta\right)\right|,\tau(\omega))e(\omega,\theta)$, where $e(\omega,\theta)$ is effort by the division's

lower-tier manager. We assume that

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

Hence, there is no intrinsic benefit to having type $\tau = 0$ or $\tau = 1$, as long as the manager gets his preferred design.

Adding the value upper-tier managers attach to the payoffs of other managers to his direct payoff, we can write the overall payoff of an upper-tier manager as:

$$u(\,|\rho(\omega, \theta - \sigma(\omega, \theta)|\,, \tau(\omega)\,)e(\omega, \theta) + \tag{4}$$
$$\int \xi(\tau(\varpi))\, u(\,|\rho(\varpi, \theta - \sigma(\varpi, \theta)|\,, \tau(\varpi))e(\varpi, \theta)d\varpi.$$

In this expression, $\tau(\varpi) \in \{0, 1\}$ is the type of upper-tier manager in division $\varpi \neq \omega$ and

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

These weights represent an "esprit de corps," by type, in the organization – i.e., you care about your co-workers, provided they share your own type.

Each upper-tier manager observes the local state $\sigma(\omega, \theta)$. He chooses $\rho(\omega, \theta)$ such that $|\rho(\omega, \theta) - \sigma(\omega, \theta)| = \tau(\omega)$ if and only if the organization is decentralized, $o(\theta) = d$.

Lower-tier managers When entering the organization, each lower-tier manager makes an effort choice, $e \in [\underline{e}, \bar{e}]$, having observed the aggregate state θ.[3] This effort is costly, where the cost $\psi(e)$ is increasing and convex with $\psi(\underline{e}) = 0$. The latter guarantees a minimum effort of $\underline{e}$. The utility of lower-tier managers is
$l(|\rho(\omega, \theta - \sigma(\omega, \theta)|\,, \tau(\omega))e(\omega, \theta)$, which we interpret as capturing a share of the upper-tier manager's decision "rent". Hence:

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

We suppose that the lower-tier managers decide on their effort after they have learned the state θ, but before they know which upper-tier manager they are matched with. Let γ be the probability that

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

$l(|\rho(\omega,\theta) - \sigma(\omega,\theta)|, \tau(\omega)) = l$, i.e., that a lower-tier manager works for a "motivated" upper-tier manager. We can then write optimal effort as

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

which is increasing in γ. Because of the timing, all lower-tier managers in the organization choose the same level of effort.

Timing The organization evolves over infinite time, with all relevant variables indexed by t. The only state variable is μ_t which represents the organizational culture, as measured by the share of type-0 managers, the equilibrium evolution of which will be derived in the next section. 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 $\theta_t \in \{0,1\}$, and $\sigma(\omega,\theta)$ for $\omega \in [0,1]$. A new generation lower-tier managers enters
2. Lower-tier managers choose effort $e_t \in [\underline{e}, \bar{e}]$
3. The leader chooses organizational form $o \in \{c,d\}$
4. Each lower-tier manager is randomly matched with one upper-tier manager. Lower-tier managers are socialized, which determines μ_{t+1}
5. If $o(\theta_t) = c$, the leader chooses a single value of $\rho(\omega,\theta_t) \in \{0,1\}$, binding for all ω
6. If $o(\theta_t) = d$, upper-tier managers in each division choose $\rho_t(\omega,\theta_t) \in \{0,1\}$
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 the organizational equilibrium of the model in a given period t with a fixed organizational culture – i.e., a fraction μ_t of type

0 managers. This allows us to map organizational culture into organization design. The results are summarized in Proposition 1. Next, we study how fraction μ_t evolves through a dynamic process, which maps the outcomes under different designs into (the change in) organizational culture. The result of these dynamics are summarized in Proposition 2. Finally, we draw out four lessons from the results in these two propositions.

5.1 Organization Design for a Given Culture

How are $e(\omega, \theta)$ and $\rho(\omega, \theta)$ determined? This will depend on whether the organization is centralized or nor. Given the timing of decisions, 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 (the Lemma and all Propositions are proven in the Appendix).

Lemma 1 *With centralization a 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 mission. However, the meaning of $\rho(\omega, \theta)$ – in terms of what the firm is actually doing – can be quite different in the $\theta = 1$ and $\theta = 0$ states.

Now let $\nu(\mu) = \mu\alpha + (1 - \mu)(1 - \alpha)$. With this notation and $\theta = 0$, $\gamma(0) = \nu(\mu)$. This is the case, because among the μ divisions with type $\tau = 0$ managers a fraction α have positive payoffs for their manager, while among the $1 - \mu$ divisions with $\tau = 1$ managers $1 - \alpha$ have positive payoffs. As a result, the ex ante probability of positive rents to lower-tier managers is $\nu(\mu)$ and effort is given by $e^*(\nu(\mu))$. Correspondingly, if $\theta = 1$ then $\gamma(1) = (1 - \nu(\mu))$, and effort is $e^*(1 - \nu(\mu))$.

Under centralization (rather than decentralization), the leader always benefits from coordination since $x = 1$. She may gain or lose from aligned projects, depending on the values of θ and α. However, the organization suffers a cost from low effort, but exactly how much also depends on parameters μ and α.

All in all, the leader's payoff is

$$\begin{cases} \Pi\left(\lambda, \left[\alpha\pi(0,0) + (1-\alpha)\pi(1,0)\right], e^*(\nu(\mu))\right) & \text{if } \theta = 0 \\ \Pi\left(\lambda, \left[\alpha\pi(1,1) + (1-\alpha)\pi(0,1)\right], e^*(1-\nu(\mu))\right) & \text{if } \theta = 1 \end{cases}$$

Decentralized control – stage 6 With decentralization, the μ divisions with a type $\tau = 0$ upper-tier manager will set $\rho(\omega, \theta) = \sigma(\omega, \theta)$. And the $(1-\mu)$ divisions with managers type $\tau = 1$ managers will set $\rho(\omega, \theta) = 1 - \sigma(\omega, \theta)$. There is always a loss from coordination since

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

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

The leader's payoff is therefore

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

Centralization versus decentralization – stage 3 Given the expressions above, it is clear the leader will choose to centralize or decentralize the organization at stage 3 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. $o(0) = d$ *if and only if*
$$\mu \geq \mu_H \geq \alpha$$
2. $o(1) = d$ *if and only if*
$$\mu \leq \mu_L \leq 1 - \alpha.$$

Proposition 1 makes intuitive sense. Suppose the leader's interests are aligned with type-0 managers, as is the case when the state is $\theta = 0$. Then, she will decentralize provided that such managers make up a sufficiently large fraction of all upper-tier managers. Conversely, she will only decentralize when $\theta = 1$ provided that sufficiently many managers are of type 1. Note that when $\lambda = 0$, so that coordination is unimportant, $\mu_H = \alpha = 1 - \mu_L$.

5.2 Socialization and 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. This culture evolves

over time through socialization across generations of managers, whose types are determined once and for all when they join the firm. 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 lower-tier managers are socialized in period t. At the cost of more algebra, we could consider longer than two-period lives in the organization, such that only a fraction of upper-tier managers retire, and a corresponding fraction of lower-tier managers get promoted, in each period. This would lead to more inertia in organizational culture.

Direct mentoring and indirect socialization We assume that being randomly matched with an upper-tier manager at stage 4 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 fitness of the two types. Specifically, let $\Delta(\mu_t)$ be the expected-utility difference between having type 0 and type 1 with a share of μ_t type-0 managers in the organization.[4] Then, a lower-tier manager becomes type 0 through mentoring if:

$$\Delta(\mu_t) + \eta \geq 0,$$

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 $G(\Delta(\mu_t))$.

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 $(1 - G(\Delta(\mu_t)))\mu_t$.

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:

$$G(\Delta(\mu_t)) + (1 - G(\Delta(\mu_t)))\mu_t. \tag{5}$$

[4] We are assuming here that socialization is based on the experience of current generation upper-tier managers. Otherwise, it would be $\Delta(\mu_{t+1})$ that mattered.

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

$$\Delta(\mu_t) + \eta \leq 0.$$

Thus, $(1 - G(\Delta(\mu_t)))$ is the proportion of type-1 managers coming from such matches. The fraction $G(\Delta(\mu_t))$ 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 $G(\Delta(\mu_t))\mu_t$.

The law of motion Multiplying (5) with μ_t, $G(\Delta(\mu_t))\mu_t$ with $1-\mu_t$, and adding the resulting expressions, we can write the equation of motion for the share of type-0 managers as

$$\begin{aligned}\mu_{t+1} &= \mu_t\left[G(\Delta(\mu_t)) + (1 - G(\Delta(\mu_t)))\mu_t\right] + (1-\mu_t)G(\Delta(\mu_t))\mu_t \\ &= \mu_t + (1-\mu_t)\mu_t 2\left[G(\Delta(\mu_t)) - \frac{1}{2}\right]. \qquad (6)\end{aligned}$$

By (6), there are three possible-steady states for the organization's culture: fully type-1 with $\hat{\mu}=0$, fully type-0 with $\hat{\mu}=1$, and interior with $\Delta(\hat{\mu})=0$ (implying $G(\Delta(\hat{\mu})) = \frac{1}{2}$). Which of these occurs occurs depends critically on the properties of $\Delta(\mu)$

Relative fitness Given (4), we can write the expected payoff difference between type $\tau=0$ and type $\tau=1$, given θ, and associated $\{\gamma(\theta), \rho(\omega,\theta)\}$ as

$$\int \{(1+\xi\mu)\, u(|\rho(\varpi,\theta) - \sigma(\varpi,\theta)|, 0) -$$
$$[1+(1-\mu)\xi] u(|\rho(\varpi,\theta) - \sigma(\varpi,\theta)|, 1)\} d\varpi e^*(\gamma(\theta))$$

In this expression, we have used the observation that γ does not depend upon in ω. We now take expectations over different realizations of θ, given μ, and the implied equilibrium choices according to Proposition 1, recalling that that $\gamma(\theta) \in \{\nu(\mu), 1-\nu(\mu), 1\}$ in accordance with the model as laid out above.

If $\theta = 0$ then the leader centralizes and sets $\rho = 0$ for $\mu \leq \mu_H$, and relative fitness of tribalists with $\tau = 0$ vs $\tau = 1$ is

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

an expression which is increasing in μ. If $\theta = 1$ then the leader centralizes and sets $\rho = 0$ if $\mu \geq \mu_L$, and relative fitness becomes:

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

an expression which has an ambiguous slope in μ.

In the complementary cases, there is decentralization and relative fitness is

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

which is clearly increasing in μ and positive (negative) whenever $\mu \geq 1/2$ ($\mu \leq 1/2$).

Dynamic paths Putting these pieces together, we can summarize the overall expression for the expected-utility difference between being a type-0 manager and a type-1 manager:

$$\Delta(\mu) = \begin{cases} \beta\hat{\delta}(\mu) + (1-\beta)\delta_L(\mu) & \text{if } \mu > \mu_H \\ \beta\delta_H(\mu) + (1-\beta)\delta_L(\mu) & \text{if } \mu \in [\mu_L, \mu_H] \\ \beta\delta_H(\mu) + (1-\beta)\hat{\delta}(\mu) & \text{if } \mu < \mu_L. \end{cases} \tag{7}$$

We make the following assumption:

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

This will hold if effort is not too responsive over the relevant range. With that assumption, we obtain (see Appendix).

Lemma 2 *If Assumption 1 holds, then for all* $\{\mu, \beta\} \in [0,1] \times [0,1]$, *there exists* $\hat{\xi}$ *such that* $\Delta(\mu)$ *is increasing in* μ *for all values of* $\xi \geq \hat{\xi}$.

Lemma 2 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 a globally increasing function. This

says that if cultural identity is strong enough – in the specific sense that tribe members put enough weight on their co-workers' payoff – we have a dynamic complementarity in the evolution of an organization's culture. The condition requires a *conformity effect*: as new members are added to the tribe, the expected payoffs to these new members are internalized strongly enough that the relative fitness of the tribe goes up. This complementarity will lead to divergent dynamics, which eventually drive organizational culture to a corner at $\mu = 0$ or $\mu = 1$.

To proceed further, we now define a critical value of organizational culture, $\tilde{\mu}(\beta)$, at which $\Delta(\mu) = 0$ in the intermediate region in (7):[5]

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

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

Proposition 2 *Under Assumption 1 and a high enough value of ξ, there are three cases*

1. *If β is close enough to one, 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 zero, 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 of Proposition 2, the organization's culture in the long run is compatible with the more frequent aggregate state. In Case 3, an intermediate range for β supports multiple stable steady states. However, given a specific initial condition for μ (and a specific value of β), the dynamics are still unique.

[5] This is the value of μ at which

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

To illustrate, consider one particular constellation of parameters with $\alpha = 1$ (so the organization is always centralized, by Proposition 1). Then, a type-0 culture is favored by higher β so that type-0 managers more often get their favored outcome. If $\beta = 1/2$, then $\tilde{\mu}\left(\frac{1}{2}\right) = 1/2$. But because of the conformity effect discussed above, a type 0 culture can emerge even if $\beta < 1/2$, provided that the initial value of μ is large enough. This, in turn, is more likely if ξ is high.

5.3 Insights from the model

Having established the main two results, we discuss four general insights that can be drawn from Propositions 1 and 2. These concern four questions: (i) how do organizational cultures and organizational designs interact? (ii) can different organizational cultures coexist under the same fundamentals? (iii) may dysfunctional cultures can survive in the long run? and (iv) when do sticky organizational cultures lead to inertia in adapting to a changing environment?

(i) Organizational culture and organizational form How does the evolution of organizational culture, μ, interact with the choice of organizational form (centralized authority versus delegation)? Propositions 1 and 2 say that there is no deterministic relation between the two. But when β is high enough for Case 1, the organization is characterized by a steadily increasing type-0 culture, together with a decentralized organization 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 observe centralization most of the time. In Case 3, when β is in an intermediate range, either of these long-term outcomes can occur depending on the initial condition.

(ii) Coexistence of different cultures Similar parameter configurations can lead to divergent paths for organizations, depending on their initial conditions. To be precise, suppose two or more organizations engage in the same activity. That is, they share the same parameters $\{\beta, \lambda, u, l, \xi\}$, and the same technologies $e^*(\gamma)$ and Π. However, Propositions 1 and 2 imply that these organizations may end up with different cultures, i.e., different long-run values of μ. In particular, this is true if parameter β lies in the intermediate range

identified in Case 3 of Proposition 2. If two organizations have different initial values μ_0 that lie on opposite sides of $\tilde{\mu}(\beta)$, we will observe two coexisting organizations, one with a type-0 and another with a type-1 long-run culture.

While this is an interesting observation, our analytical framework so far does not allow for interactions between different organizations. Since firms, bureaucracies and political parties typically do not operate in a vacuum, this is an important omission. In the next section, we study different applications of the theory, which illustrate different implications of our general framework. Some of these applications do allow for organizational interactions. 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. In order to obtain a sharper result, we assume that the leader's payoff satisfies (3), which holds in most our applications in Section 6. For this case, we have the following simple comparison:

Proposition 3 *If the leader's payoffs are multiplicative, her payoff for* $\mu = 1$ *is greater or smaller than for* $\mu = 0$ *depending on whether*

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

Note that as $\beta \to 1$, then $\mu = 1$ is the long-run outcome and according to Proposition 2 a type-0 culture emerges. Similarly as $\beta \to 0$, then $\mu = 0$ is the long-run outcome and a type-1 culture is 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 we can have multiple steady states depending on the initial condition. Then, there is no guarantee that the organization will converge to a culture that maximizes long-run payoffs. Indeed for β close to $\frac{1}{2}$, we can make the gain to the leader from having her preferred long-run culture arbitrarily large by varying $\pi(0,0)$ compared to $\pi(1,1)$. Hence, highly dysfunctional cultures can emerge in the long run.

The result that organizational cultures can be dysfunctional is perhaps not too surprising, given that cultural evolution hinges on $\Delta(\mu)$, which reflects the expected payoffs for managers. These, in turn depend on the design

choices by leaders, which are optimized period by period. However, they do not depend directly on the leader's payoffs.

This aspect of our model ties in with earlier discussions around the Coase theorem whether organizational forms are efficient. There is also a parallel with the argument in Acemoglu (2003) that lack of commitment by current decision-makers is a barrier to efficiency. The culture that emerges in long-run equilibrium is dysfunctional from the leader's viewpoint. If the owners could take a long-run view and delegate long-run control of the organization to a leader with a particular tribal preference, they would prefer to do so.[6]

(iv) Organizational inertia Another upshot from the model is that the culture is immune to change even if some parameter values are permanently altered. Thus organizational cultures can limit adaptability, as in the IBM-example discussed in Section 3. To illustrate this, consider two values of $\beta \in \{\beta_L, \beta_H\}$ and assume 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 $\beta = \beta_H$ (to see this use the condition in (7) 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 (7) and Proposition 2, for all β such that

$$\beta\xi e^*(1) + (1-\beta)\, e^*(1-\alpha)\,[1 - 2\alpha + (1-\alpha)\xi] > 0 \tag{8}$$

the culture persists at the point $\mu = 1$.[7] This is because $\Delta(1) > 0$. From (8), we see that there exists a critical value of β, given by

$$\hat{\beta}_L = \frac{e^*(1-\alpha)\,[2\alpha - 1 - (1-\alpha)\xi]}{\xi e^*(1) + e^*(1-\alpha)\,[2\alpha - 1 - (1-\alpha)\xi]}, \tag{9}$$

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

[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.

[7] This implies that

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

This result suggests that a significant shift in the environment may be needed to start a process of cultural change. Note also that $\hat{\beta}_L$ is decreasing in ξ, the loyalty bond in entailed in organizational identity. The friction created by culture is thus greater the stronger these bonds. Thus, our model demonstrates how identity-based cultures naturally inhibit organizational adaptability – i.e., such cultures promote organizational hysteresis. This conforms to frequent claims in the literature on organizational behavior.

6 Applications

In this section, we put the model to work in four specific applications to show how it can illuminate questions around the role of culture in organizations such as bureaucracies, firms, and political parties. The specificity of the approach can also generate new insights, which may merit further development 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 which use similar technologies and have similar access to resources. Classic accounts of public bureaucracy, such as Wilson (1989), emphasize culture and values as elements that can explain inertia and resistance to change. Because the scope to use traditional tools of performance management is restricted – as key public-sector service outputs are hard to measure, formal incentive arrangements have limited bite – good service delivery may have to rely on the underlying 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 measurement tools and insights from their analysis of private firms, Bloom et al (2014, 2015) find the same differences in bureaucratic management as in private management, and that management styles are 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 and quality improvement services 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 running 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 be looking at future fund-raising possibilities or higher student-tuition fees. However, faculty members may have their own priorities, preferring to engage in teaching or research with their productivity being higher when they are performing the task they value the most. Moreover, transmission of those values from senior to junior faculty may be an important part of the cultural transmission process.

Application of the model In running any kind of bureaucracy, a key issue is how much to centralize aspects of the mission choice and how much to allow local discretion in decision-making. Leaders may be concerned that decentralization may lead the organization astray from its main priorities. Our framework is useful for understanding the challenges of building an effective organizational culture, which best serves the ultimate beneficiaries such as victims of crime, patients, or students.

In our application to bureaucracies, we thus interpret ω as reflecting different providers in a system of police precincts, hospitals or schools. The choices $\rho(\omega, \theta)$ would then represent aspects of the mission: where to orient resources to fight crime, which medical treatments to prioritize, or what kind of school curriculum to develop. The variable $\theta \in \{0, 1\}$ is a reflection of where the leader believes that the organization's priorities should lie, while $\sigma(\omega, \theta)$ allows local variation in the mission to reflect local priorities.

In these applications, lower-tier managers are best thought about as professionals who deliver services and from whom the senior management are drawn. In practice, not every front-line professional becomes a senior manager, but it is common to appoint senior public managers among previous practitioners – school principals are often former teachers. These managers

are more motivated when they get to undertake the activity that they like most.

Our core framework for understanding organizational culture can now be applied in this context and we simplify the organizational objective of the leader to be of the form in (3), i.e.,

$$\Pi\left(\lambda\left(2x-1\right)^{2},\int_{0}^{1}\pi\left(\left|\rho\left(\omega,\theta\right)-\sigma\left(\omega,\theta\right)\right|,\theta\right)d\omega,e\right)= \tag{10}$$

$$\hat{\phi}\left(x\right)\times e\times\int\pi\left(\left|\rho\left(\omega,\theta\right)-\sigma\left(\omega,\theta\right)\right|,\theta\right)d\omega,$$

where $\hat{\phi}\left(x\right)=\frac{1+\lambda\left[2x-1\right]^{2}}{1+\lambda}$ represents possible spillovers across service providers from coordinating their strategies. All the results in Propositions 1-3 apply straightforwardly in this case.

We will further simplify by studying a symmetric case where

$$\pi\left(1,1\right)=\pi\left(0,0\right)=\pi_{H}>\pi\left(0,1\right)=\pi\left(1,0\right)=\pi_{L}. \tag{11}$$

This implies that there is no intrinsic advantage to either of the organization's possible priorities.[8] The assumption in (11) also implies

$$\int\pi\left(\left|\rho\left(\omega,\theta\right)-\sigma\left(\omega,\theta\right)\right|,\theta\right)d\omega=\pi_{L}+\left(\pi_{H}-\pi_{L}\right)\left[\theta+\mu-2\mu\theta\right]$$

in the decentralized case.

We now discuss how the model can cast light on three frequently discussed features of public bureaucracies: (i) dilemmas of top-down control, (ii) heterogeneous performance, which cannot be explained by resources or technologies, and (iii) institutional inertia and resistance to reform.

Dilemmas of top-down control Deciding 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). A frequently made claim is that decentralization works best to take advantage

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

$$\left(1-2\beta\right)\left[\frac{e^{*}\left(1\right)-\alpha e^{*}\left(1-\alpha\right)}{\left(1-\alpha\right)e^{*}\left(1-\alpha\right)}\right]\lessgtr\left(1-2\beta\right)\frac{\pi_{L}}{\pi_{H}}.$$

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. Our result in Proposition 2 shows that culture will support the objectives of the center in the long run when goals are clearly defined, in the sense that 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 that there will be top-down control when the centre and delivery units are poorly aligned. Specifically, Proposition 2 shows that if β is close to 0 or 1, such clashing interests are unusual and that this raises organizational efficiency. However, organizations where β is close to $\frac{1}{2}$ are more likely to see regular conflict between leaders and management and a resulting inefficiency due to lower managerial effort. So mission clarity is associated with better performance, a theme in Wilson (1989).

Heterogeneous performance Our model speaks straightforwardly to the central puzzle that bureaucratic performance differs in apparently similar organizations (in terms of technology and funding) due to emergent cultures. This is straightforward to see by applying case 3 of Proposition 2, which shows that close to $\tilde{\mu}(\beta)$ organizations may follow different paths. If the state θ is common across organizations, then at a point in time when $\theta = 0$, organizations with a culture approaching $\mu = 1$ will perform better than those with μ approaching 0. The converse will be true when $\theta = 1$.

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 the case of academia). To see how this can occur 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 entrenched values among managers. The organization can try to deal with this by centralizing, but this will result in an efficiency loss due to effort being lower at $e^*(1-\alpha)$ under centralization rather than at $e^*(1)$ 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 the model. In future work, it will be interesting to consider the role of hiring and firing on such a transition path.

6.2 Firms, Productivity, and Corporate Cultures

Organizationally, there is no fundamental difference between a public bureaucracy and a management bureaucracy in a firm. Thus the same insights as in the previous subsection largely carry over to this context. However, a private firm may be subject to a harder budget constraint as it has to survive in the market. Our framework allows us to think about these issues. We thus develop a model that can capture the heterogeneity of firm productivity found in the data and the possible links to different management styles, as those studied by Bloom and van Reenen with different coauthors. A simple way to do so is to use a "span of control" model as in Lucas (1978), where managers in each division of the firm can hire workers and the firm's leader is a CEO who cares about profits.

Technology Suppose that the productivity level of a typical division in one of these firms 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 $\hat{\phi}\left(x\right)=\frac{1+\lambda[2x-1]^2}{1+\lambda}$ reflects the value of coordination for productivity.[9] We will maintain, the case where $\pi(\left|\rho\left(\omega,\theta\right)-\sigma\left(\omega,\theta\right)\right|,\theta)$ satisfies (11). Independently of the organization of the firm, each division can hire labor $l\left(\omega\right)$ with a divisional decreasing-returns production function: $\nu^{1-\zeta}l^{\zeta}$ where $\zeta<1$. Laborers l can be freely hired at wage w.

We can now think about how corporate culture might affect a firm's management style – embodied in $\rho\left(\omega,\theta\right)$ – which, in turn, shapes organization design. The latter choice can become a source of competitive advantage by affecting the firm's profitability, which depends on the culture as embodied in μ. Parameter β now captures how different types of divisional management

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

decisions shape firm productivity. The aggregate shock θ can be thought of as different states of the world, where different management activities are more or less productive for the firm. A culture clash arises when upper-tier managers have a proclivity towards activities which are not the most productive for the organization. In the first instance, this will generate something similar to the public-service application in Subsection 6.1.

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

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

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

where $\hat{\zeta} \left(w \right) = \left(\frac{w}{\zeta} \right)^{-\frac{\zeta}{1-\zeta}}$. In this setting, division-level and firm-level heterogeneities depend on organization and project decisions by upper-tier managers – think about the latter as the firm's "management style". In this sense, the model in this section provides a theoretical micro-foundation for the empirical analysis in Bloom and Van Reenen (2007) and Bloom et al (2012).

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

$$\begin{aligned} & \Pi \left(\lambda \left(2x - 1 \right)^2, \left[\int_0^1 \pi(\left| \rho \left(\omega, \theta \right) - \sigma \left(\omega, \theta \right) \right|, \theta) d\omega \right], e \right) \qquad (12) \\ = \; & \left(1 - \zeta \right) \hat{\zeta} \left(w \right) p^{\frac{1}{1-\zeta}} \hat{\phi} \left(x \right) \left[\pi_L + \left(\pi_H - \pi_L \right) \left[\theta + \mu - 2\mu\theta \right] \right] e. \end{aligned}$$

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

Centralized control, management form, and firm heterogeneity This application fits the framework of our core model so that Propositions 1-3 all apply. The model therefore provides some insights into possible theoretical foundations for Bloom et al (2012), who uncover a positive correlation between decentralization in firms and their performance. However, our model predicts that decisions to decentralize, management culture, managerial effort, and firm performance are all jointly determined. Thus there is a complex

web of causal interdependencies 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 senior 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 provides a foundation for the existence of persistent heterogeneities in productivity and profits among firms, even when the same market conditions and technologies are available to all of them. Firms that evolve better cultures will be more productive and profitable. Our framework suggests that homogeneity is only like to emerge when β is close to zero or one – i.e., when the environment is highly predictable and when it supports one specific type of organizational culture. When the organization 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 above, a key difference between public services and private firms is that market discipline can provide a bound on cultural inefficiencies in the latter case. We now explore this idea, focusing on the case where $\lambda = 0$ – i.e., we abstract from coordination gains or costs. Suppose that, to stay in business, each firm has to pay a fixed cost F in terms of labor in each period and that this cost is paid prior to θ being realized. The existence of a fixed cost allows for the possibility that a market-selection mechanism works, whereby in some circumstances only firms with certain kinds of cultures can 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 $\theta = 1$. Each type of culture suffers an efficiency loss in states where managers have to act against their preferences. But cultures may also motivate managers and enhance effort.

Can both type-0 and type-1 cultures coexist, or does the market constraint make one of them infeasible? To probe this question, suppose that 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. Note that, 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*

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

This bounds the allowable inefficiency among firms with different long-run cultures. Via the LHS of the inequality in Proposition 4, this bound depends on the predictability of the aggregate environment, β, the correlation across local conditions, α, and the size of 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 in low-wage settings with high prices so that profits are high, or when the fixed costs are low. All of these contribute to a weak market test.

If there is coexistence, one of the cultures 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 as observed in the focus of management on particular problems and their preferences to tackle them in particular ways.

The flip side of Proposition 4 is the case where market conditions weed out one of the cultures. In the inequality fails, the market test will weed out one of the cultures in the long run. Unsurprisingly, a hard budget constraint does reduce the long-run permissible degree of cultural inefficiency. Shifts in market conditions – like deregulation or opening up to trade (which could lower p or raise w) – could thus contribute to eliminating inefficient cultures.[10]

[10] Our model also predicts that the aggregate distribution of corporate cultures in a

6.3 Culture and Management Focus in IBM

Our model can also be used to revisit the IBM case study. In Section 3, we emphasized the challenge for this organization to adapt its culture to a new product line. To make this point more precise, suppose the firm can specialize in one of two products: mainframes, M and PCs, P. Let $\pi_M(\theta)$ and $\pi_P(\theta)$ be the profits associated with the two depending on market conditions as summarized by aggregate state θ. Also, assume that $\pi_M(0) > \pi_P(0)$ and $\pi_M(1) < \pi_P(1)$.[11] Finally, interpret managers with $\tau(\omega) = 0$ as adopting a mainframe-oriented culture and those with $\tau(\omega) = 1$ as adopting a PC-oriented culture. Hence, they tend to focus on missions which enhance the technology of the products they identify with.

Suppose further that μ is the proportion of management focused on activities that enhance mainframe technologies. With the multiplicative performance function in (3), we can write the firm's profits (the leader's payoff) as

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

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

Consider a firm as the old IBM, where $\mu = \beta = 1$, as the result of a cultural-convergence process in Proposition 2. As the state is always $\theta = 0$, this firm will operate on a decentralized basis where all lower-tier managers are motivated and put in effort $e^*(1)$. Moreover, the uniform culture and work habits will give full coordination on mainframes with $y(0) = x = 1$. Profits are therefore given by $\lambda[\pi_M(0)]e^*(1)$.

---

market will affect the equilibrium price 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 $Q(\theta)$ being the total industry output in state θ and suppose that θ is common to all firms. Suppose that $\Omega(\theta)$ is the proportion of firms which have evolved a culture where the management is aligned with the firm when the state is 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 then lower in states of the world which 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 $\pi_M(0) = \pi(0,0)$

A change in market conditions What happens if β falls, making state $\theta = 1$ more common as PCs becomes more attractive relative to mainframes? In state $\theta = 1$, the natural response of the management is to centralize the organization and choose to impose the same PC-oriented projects on all divisions, since $\pi_M(1) < [\pi_P(1)\alpha + (1-\alpha)\pi_M(1)]$. As local information is lost, this will lead to some advances in PCs and some in mainframes by the "misdirected" managers. Profits are now given by $\lambda[\pi_P(1)\alpha + (1-\alpha)\pi_M(1)]e^*(1-\alpha)$.

These profits will be lower than the profits of a firm with a PC culture, $\mu = 0$, which will be $\lambda[\pi_P(1)]e^*(1)$. Such a firm elicits effort $e^*(1)$ from its managers, and can decentralize projects to get better aligned decisions with profits $\pi_P(1)$ for all divisions. On both counts, 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 its managers perceive the change in β. If the new environment is one with $\beta = 0$, a process of cultural change will begin. But during the transition, IBM will have to wait for sufficiently many managers to turn over in the dynamic socialization process. Following the analysis in Section 5, however, if the "death of the mainframe" is still in doubt – such that β is higher than $\hat{\beta}_L$ defined in (9) – culture may not change. This is especially likely with a strong *esprit de corps* among the managers (high ξ).

This analysis illustrates not only the narrative of IBM and its slow adaptability due to a strong culture. It also allows us to articulate similar concerns, which are now being expressed about the prospects for Google, as it tries to adapt to greater competition and new product lines. For example, taking on Facebook and adapting to the use of mobile apps has created key challenges.

Strong organizational cultures can be very powerful in stable environments, but create inertia when there is a need to adapt. Then, the strong culture risks becoming dysfunctional. It would be interesting in further work to combine this insight with the insights from the analysis in Subsection 6.2 of a market-selection process. We conjecture that the market may eventually weed out "dinosaur" cultures, but such weedouts may be slower in markets with less competition.

6.4 Political Parties

Finally, we show how to 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 dimension of political parties has not been studied a great deal, although standard political-science treatments of parties do point out that centralized authority is sometimes needed but can also be too strong (Cox and McCubbins, 2003). It is nevertheless important. For example, Willis et al (1999) argue convincingly that the differential structure of Latin American parties – e.g., very centralized parties in Mexico and decentralized parties in Brazil – are important to understand the differential decentralization of political powers on the continent.

Voter preferences Consider a set-up with two parties $P = A, B$. Each of these parties has a leader who manages a multi-division organization – with local party heads and party workers, analogous to the upper-tier and lower-tier managers – like the one studied in Sections 4 and 5.

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

$$W(|\rho(\omega,\theta) - \sigma(\omega,\theta)|, x, \theta, e) = \lambda(2x-1)^2 + \pi(|\rho(\omega,\theta) - \sigma(\omega,\theta)|, \theta) \cdot e. \tag{13}$$

The first term represents a national policy which has higher value to the voters the more coordinated the actions by the party and λ indexes the importance of this national issue. The second term captures a policy targeted to district ω, which is magnified by the effort e local party workers put into policy design. Furthermore, voters get an extra χ of utility under party-B rule. In other words, χ is a popularity shock in favor of party B, and is continuously distributed with mean zero, $E(\chi) = 0$, and a symmetric single-peaked density. By symmetry, the c.d.f. of the popularity shock, Γ has $\Gamma(0) = 1/2$. The χ-shock is realized after policy-design choices at stage 5 or 6, but before the election that occurs in each period. Voter preferences accord with those of type-1 district leaders, when $\theta = 1$, which occurs with probability $1 - \beta$, and with those of type- 0 district leaders when $\theta = 0$, which occurs with probability β. Again we work with (11) but with $\pi_H = 1$ and $\pi_L = 0$.

Winning probabilities When the parties offer policies $\{\rho^P(\omega,\theta), x_P, e_P\}$, then voters in district ω will thus vote for party A if

$$\chi \leq W(\left|\rho^A(\omega,\theta) - \sigma(\omega,\theta)\right|, x_A, \theta, e_A) - W(\left|\rho^B(\omega,\theta) - \sigma(\omega,\theta)\right|, x_B, \theta, e_B).$$

Observe also that

$$\int W(\left|\rho^P(\omega,\theta)-\sigma(\omega,\theta)\right|,x_P,\theta,e_P)d\omega$$
$$=\lambda(2x_P-1)^2+e_P\cdot[(1-\theta)x_P+\theta(1-x_P)]\equiv w(x_P,e_P:\theta)$$

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

$$\begin{aligned} p(\mu_A,e_A,\theta;\mu_B,e_B) &= \\ \text{Prob}[\chi &\leq w(\mu_A,e_A:\theta)-w(\mu_B,e_B:\theta)] \\ &= \Gamma\left(w(\mu_A,e_A:\theta)-w(\mu_B,e_B:\theta)\right). \end{aligned} \tag{14}$$

The probability of winning for party B is just given by $1-p(\mu_A,e_A,\theta;x_B,e_B)$.

Substituting from (13) into (14), we see that the payoff of a party is

$$\Pi\left(\lambda(2x-1)^2,\int_0^1\pi(|\rho(\omega,\theta)-\sigma(\omega,\theta)|,\theta)\,d\omega,e\right)= \tag{15}$$
$$\Gamma\left(\lambda(2x-1)^2+e\cdot[(1-\theta)\mu+\theta(1-\mu)]-\text{constant}\right)$$

which fits the general model although not (3). Propositions 1-2 still apply.

The constant in (15) depends on the strategy of the other party. However, it is clear that the party's objective is equivalent to maximizing $\lambda(2x-1)^2+e\cdot[(1-\theta)\mu+\theta(1-\mu)]$. This amounts to each party choosing a party organization in each period depending on θ. Under centralization, the policy vector is chosen centrally, while with decentralization it is delegated to local party district managers.

Decentralization and party cultures The analysis hints at a novel aspect of electoral competition, which has not received much attention in the academic literature to date. For example, Green parties in European countries like Germany and Sweden started out as very decentralized organizations accommodating a strong party culture among engaged local party workers. As these parties gradually came to take part in national and regional coalition governments, party leaders saw a need to centralize policy-making – think about this as a higher weight λ on coordinated policies in the model. But this was met with complaints among party members and former party leaders. Our model can be used to think about such developments as rational responses to changes in the environment as perceived by party leaders.

The analysis adds the insight that such a change in party objectives and party organization would gradually change the prevailing party culture. 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 its 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, it follows from Proposition 2 that – in the long run – party A will evolve a different party culture with $\mu^A = 1$, compared to party B where $\mu^B = 0$. In our model, these identities will be associated with loyalties among party workers. Both party cultures can coexist and, as we see below, one party could spend more time in office even if the party fundamentals are similar, simply on the back of their party structure being different. In the long run, parties may or may not be decentralized, depending on the value of λ, i.e., to what extent greater coordination is valuable to winning. Studying this further in specific party contexts would be interesting.

A competitive cultural advantage? We now investigate how a party culture can become an electoral asset or a liability. Consider the case where $\mu^A = 1$ and $\mu^B = 0$. The winning probability for party A is then $p^A = \Gamma(w^{A*} - w^{B*})$, 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 $\Gamma(w^{A*} - w^{B*}) \gtreqless 1/2$. Under these conditions, we have

Proposition 5 *Suppose that* $\mu_L > 0$ and $\mu_H < 1$ *and that 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* :

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

Suppose that $\theta = 0$. If both parties decentralize then $x = \alpha$ for both. However, $\pi = e^*(1)$ for A and $\pi = 0$ for B. Thus, for B to be able to compete with A, B must centralize. 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.

This advantage comes primarily 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 and there is alignment between the centre and the local party managers. Since the same θ shock hits party B, 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 centre is pushing against what the senior party managers want. While B also 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, whether a party culture is suitable for winning elections is context specific. That is, in the short run it depends on the realization of θ, and in the long run average electoral success depends on β. Differences in political advantage due to party culture will be large if there is stronger political competition represented by a density function for the popularity shock χ which is larger around its mean (zero). 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]

Using some of the core insights of the model, we can also explain the difficulty of adapting party cultures to changed political circumstances following a permanent shift in β which favors one party. Even though it may be in the interest of one party to modify its party culture, this may be difficult for reasons which we explored above, giving it a lasting electoral disadvantage.

7 Final Remarks

We have proposed a model of organizational culture where identity-based socialization of managers leads to cultural dynamics. The framework generates a range of insights on the interplay between organizational culture and organization design with implications for performance. The model makes precise conditions under which different organizational cultures emerge in the long run. Whether the organization is centralized or decentralized is endogenous and depends on internal conflicts of interest, which reflect tensions between the organization's culture and the leader's state-dependent objectives. We also propose four specific applications of these general ideas.

[12] To see this concretely suppose that ξ 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.

The framework could be developed in various ways. Hirschman (1970) famously emphasized three sources of organizational dynamics: exit, voice and loyalty. We focus on the role of loyalty, as transmitted by social identity. But the model could be extended to include exit and voice. Exit would reflect that organizations under stress often hire managers from the outside to by-pass those who have become socialized into particular modes of behavior. It would be interesting to consider this in future research by embedding organizations in a market for managers. Voice would reflect managers having a more direct say in the centralized operation of the organization. For example, allowing senior managers to vote over the mission – e.g., the ρ chosen under centralization – would give an advantage to the majority culture. But one could study a variety of voice mechanisms, including the way leaders are selected and how much say insiders have in that process.

A wider set of issues about governance and leadership could be explored with our framework. For example, a leader allowed to pursue a particular organizational objective could have a long-run transformational effect. But she may also create short-run unhappiness, by demotivating existing managers, as she attempts to transform the culture. The way leaders are evaluated will then be important – e.g., whether poor short-term performance is tolerated and not interpreted as the result of leader 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.

A richer theory of what leaders do would also be interesting. We have confined their role to changing the authority structure. However, as Weber (1922) emphasized in his theory of charismatic leadership, inspiring leaders can serve as catalysts for cultural change, quite apart from the sticks and carrots at their disposal. This would somehow allow the leader to have a more direct effect on μ_t in our model.

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 thus study how the evolution of democratic values interacts with reforms of democratic institutions. Research on the interplay between formal rules and cultural values remain scarce – further explorations will make 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: Princeton University Press, 2010.

[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,"

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

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

[14] 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.

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

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

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

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

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

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

[21] 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

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

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

[24] 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.

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

[26] 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.

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

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

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

[30] 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.

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

[32] 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.

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

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

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

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

[37] 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.

[38] 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.

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

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

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

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

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

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

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

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

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

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

[49] 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.

[50] Whyte, William H. [1956], *The Organization Man*, Simon & Schuster.

[51] 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.

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

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

Appendix

A 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

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

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 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:

$$\xi \geq \max\left\{\frac{e^*(1-\nu(\mu_L))[2\alpha-1]}{[1-2\mu_L]e^*(1)-e^*(1-\nu(\mu_L))(\alpha-\mu_L)]}, \frac{e^*(\nu(\mu_H))[2\alpha-1]}{e^*(1)[2\mu_H-1]-e^*(\nu(\mu_H))(\mu_H+\alpha-1)]}\right\} \tag{16}$$

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

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

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

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

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

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

With $\mu = 0$ it is instead

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

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

$$\begin{aligned}&\beta\left[\Pi\left(\lambda, \pi\left(0,0\right), e^*\left(1\right)\right)-\Pi\left(\lambda, \left[\alpha\pi\left(0,0\right)+\left(1-\alpha\right)\pi\left(1,0\right)\right], e^*\left(1-\alpha\right)\right)\right] > (<) \\ &\left(1-\beta\right)\left[\Pi\left(\lambda, \pi\left(1,1\right), e^*\left(1\right)\right)-\Pi\left(\lambda, \left[\alpha\pi\left(1,1\right)+\left(1-\alpha\right)\pi\left(0,1\right)\right], e^*\left(1-\alpha\right)\right)\right].\end{aligned} \tag{17}$$

In the multiplicative case, this boils down to

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

which yields the condition in the proposition.

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

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

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

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

Putting these together yields the result.