Abstract

The main result of this paper is that, when joint action is called for, people prefer to associate with others whose beliefs or values are similar to their own, since such others take ‘correct’ decisions from the focal agent’s perspective. This causes a homogeneity of beliefs and values within a firm or organization, which corresponds directly with the sociological definition of organizational culture as shared beliefs or values. The paper explores the consequences of such theory of corporate culture, in particular the conditions which create a ‘strong’ or highly homogenous culture. It also summarizes existing and new research on the benefits and costs of belief homogeneity, and discusses ways to manage it.

1 Introduction

The saying ‘birds of a feather flock together’ suggests that people associate with similar others. This paper explores the economic incentives for doing so and the consequences for organizations. The main result is that, when joint action is called for and people have openly differing beliefs about the right course of action, people tend to gain subjectively from associating with others who have beliefs or values that are similar to their own\(^1\). The reason is, loosely speaking, that such similar others make the ‘right’ decisions from the focal agent’s perspective.

One important implication is that firms and organizations will be more homogenous in their beliefs and values than society at large. Such homogeneity corresponds to the leading sociological definitions of corporate culture as shared beliefs, values, or assumptions (Schein 1985). Corporate culture is a central concept in the management and organizational literature. It is difficult to find a discussion of mergers or corporate change that does not mention corporate culture. More in general, culture has been credited with the success and blamed for the demise of some major corporations. Despite its importance, the topic has received only limited attention in the economic literature. (A discussion of the relevant literature follows later.) The contribution of this paper is to present a theory that starts from economic micro-foundations, is able to account for both the formation and the effects of corporate culture, and is in line with the organizational literature. The

\(^1\)I will also consider when it might be optimal to explicitly associate with others with different beliefs or values.
The paper then applies the theory to study the determinants of the strength of corporate culture. It also summarizes recent research on the performance implications of such homogenous beliefs.

The basic premises of the theory are simple. On the one hand, I assume that an employee’s payoff depends on his firm’s performance. Employees of successful firms typically share in the rents via lower risks of lay-off and higher likelihoods of promotion and wage increases. On the other hand, the firm’s performance depends on its employees’ decisions, but employees may openly disagree, in the sense of differing priors\(^2\), about the optimal course of action.

The paper then first considers whether or when it is optimal to associate with similar others. It obtains essentially three conclusions. First, absent incontractible effort, it is always (subjectively) optimal to have colleagues whose beliefs are identical to yours. Second, however, this does not imply that it is always better to work with people whose beliefs are more similar to yours\(^3\), due to the interactions between different agents’ actions. Nevertheless I show that in many important situations, this local comparative statics still holds, at least in expectation. Third, it may be optimal for agents to work with others who expressly have differing beliefs when incontractible effort, especially information collection, is important. Overall, this part concludes that, with some important exceptions, rational agents will want to associate with similar others. This creates a homogeneity of beliefs within firms.

The paper then links this to the theory of corporate culture. In particular, the leading definition of corporate culture in the sociology and management literature is that of culture as shared beliefs or values (Schein 1985). Using both static and dynamic models, I then analyze the formation of such culture and derive comparative statics. One of the key insights is the strong path dependence of culture, both in terms of content as in terms of strength. This explains one of the key stylized facts in the literature: that culture depends strongly on the founder or early leaders of an organization and that their influence can work long after they are gone (Schein 1985, Baron et al. 1999).

Note that this theory in itself is neutral on the desirability of a strong culture. Culture is a by-product of subjective individual rationality and not some optimally chosen firm characteristic. There are nevertheless important performance implications, some of which are summarized at the end of the paper on the basis of recent research on the benefits and costs of belief homogeneity. The paper concludes with a discussion of ways in which culture might be influenced.

Before proceeding with an overview of the literature, a few remarks are in order. First of all, although most of the paper is formulated in terms of firms and in terms of differing beliefs, the results can readily be extended to organizations in general and to heterogenous values or utility functions. Second, this theory of association is more applicable to the selection of top management than to that of the rank and file. Finally, the model makes abstraction of any processes that would cause further convergence of the beliefs or values of a given set of employees, as considered for example in Lazear (1995). While such processes are important to fully understand culture, it would overload the current model. Future research should definitely deepen our understanding of both processes and their interaction.

The paper that is most closely related is Cremer’s (1993) analysis of corporate culture. He defines culture as a stock of shared knowledge and argues that it improves the efficiency of information processing. While most of the paper is an informal discussion on communication, it presents an interesting team-theoretic model that shows how shared beliefs may improve coordination. A very analogous conclusion in the context of this paper is drawn in section 8. The main weaknesses

---

\(^2\)For a discussion of this approach, see Morris (1995) or Van den Steen (2002c).

\(^3\)In other words, while your subjective utility does not necessarily always increase as the beliefs of your colleagues get more similar to yours, it will be be maximal when their beliefs become identical to yours.
of Cremer’s paper are linked to his assumption that culture is driven by incomplete information. For one thing, it is difficult to imagine that the difference in culture between Apple and GM is caused by the fact that Apple employees have shared different information than GM employees. This is closely related to Cremer’s own remark that values don’t fit well into his theory. A second implication is that the only real cost of culture in this model is the cost of communicating the information. Beyond that, more culture is always better. Dysfunctional cultures cannot exist in this view. Moreover, the theory is weak in explaining why culture is so difficult to change or how firms in the same industry with similar history might have very different cultures, given that culture is essentially a choice variable in this model.

Lazear (1995) defines culture as ‘shared beliefs, values and technology’ and considers a ‘genetic’ evolutionary model of corporate culture, built on the assumption that culture is somehow contagious. This work is complementary to the current paper in that it studies how culture might further evolve once the firm is formed. It also contains a broad overview of the non-economic literature on the topic. The strength and weakness of the paper lies in its level of abstraction. Without any micro-foundations, the paper assumes that cultural traits are somehow contagious, that management can somehow affect the transmission of culture and that homogeneity is somehow valuable. Van den Steen (2001) suggests a different mechanism how strong managerial beliefs may attract like-minded people and thus cause homogenous beliefs within an organization. While the focus of that paper is on the effects of strong managerial beliefs, or vision, its conclusions fit directly with the literature on cultural change (Kotter and Heskett 1992, Schein 1985).

There is some other work on corporate culture that is less related to the current paper. Kreps (1990) identifies culture essentially with a firm’s reputation for dealing in a specific way with unforeseen contingencies and argues that such reputation is useful to protect employees against abuses of authority. This perspective is quite different from the rest of the literature on corporate culture. Hermelin (1999) summarizes and reinterprets the existing research, but also adds to it by linking the topic with insights in other fields of economics, such as IO. In one model, for example, he assumes that culture is a fixed cost investment that lowers the variable cost and then derives industry-level implications. Rob and Zemsky (2000) present a theory in which firms differ in the stationary state levels of cooperation among their employees, which they link to the notion of corporate culture.

The discussion on the costs and benefits of belief homogeneity is related to the literature on advocacy (Milgrom 1981, Milgrom and Roberts 1986, Rotemberg and Saloner 1995, Dewatripont and Tirole 1999). While there exists a literature on clubs and associations (Ellickson, Grodal, and Scotchmer 1999), these papers simply assume that people have preferences to associate with specific others. From a methodological point of view, it is finally interesting to note that, by allowing differing priors, the first model of the paper introduces agency considerations in a team-theoretic setting (Marschak and Radner 1972).

The contribution of this paper, then, is to present a theory that starts from economic micro-foundations, is able to account for both the formation and the effects of corporate culture, and is in line with the organizational literature. The main insight is probably the view of culture as a by-product of subjective individual rationality. It is the first economic theory to really explain why cultures may be dysfunctional and, with the partial exception of Lazear’s genetic model, the first that can make sense of ‘culture as shared values’. It also has immediate implications for the performance effects of culture.

The next section studies the saying that ‘birds of a feather flock together’. Section 3 motivates the definition of corporate culture as shared values or beliefs. Section 4 discusses the issues involved
in studying the formation of corporate culture, while sections 5 through 7 consider both static and
dynamic models of culture formation. Section 8 summarizes findings on the benefits and costs of
homogenous beliefs, while section 9 discusses some ways to influence culture. Section 10, finally,
concludes.

2 Birds of a Feather ...

This section considers the agents’ subjective incentives, when undertaking a cooperative effort, to
associate with others with similar beliefs and values. Its key conclusions are that, absent incon-
tractible effort,

1. it is always optimal to work with people who have identical beliefs to your own.

2. it is often better to work with people who hold beliefs that are more similar to your own.

These conclusion, especially the first one, may seem surprising. Consider for example a situation
with 2 agents who have to simultaneously choose an action \( x_i \in \mathbb{R} \) with each getting a payoff
\( u_i = -\min_{i \in \{1,2\}} (x - x_i)^2 \), where \( x \) is the true but unknown state of the world. Clearly, no
matter what your beliefs about \( x \), it is optimal for the agents to undertake different actions, i.e. to
‘experiment’. Wouldn’t it be better then to work with an agent who has beliefs that are different
from your own? The answer is negative for the following reason. If you work with an agent who
holds exactly the same beliefs as you do, then you agree on the optimal amount and the optimal
form of experimentation and you can coordinate on that. You cannot do any better than this. A
formalization of exactly this argument will be used below to prove this proposition in full generality.
Note, however, the important qualification that there be no incontractible effort. Subsection 2.3
below discusses how such effort may make it optimal to work with people who have a different
opinion. A second issue, on which I come back in section 8, is the question whether and how the
conclusions might change when change or innovation is important.

2.1 A global result

Let there be \( K \) agents, who each have to make a decision \( x_j \in \mathbb{R} \), and let there be a true state
\( x \in \mathbb{R} \). The true state \( x \) is unknown, but all agents have subjective beliefs about it. These beliefs
may openly differ, i.e. agents may have differing priors. Assume that each agent’s payoff is a share
\( \alpha_i \) of the overall payoff \( \Pi(x_1, ..., x_j, ..., x_K, x) \). The formal approach is thus similar to Marschak
and Radner’s (1972) team theory, in that all players have essentially the same objective function,
but allows for differing beliefs. Note that this introduces agency concerns in team theory. I also
make the following assumption.

Assumption 1 \( E_i[\Pi] \) is continuous in the actions. For any set of beliefs about \( x \), there exists a
Nash equilibrium. Players coordinate on a Pareto-efficient equilibrium whenever one exists.

Proposition 1 Under A1, each agent is (subjectively) better off when all other agents have beliefs
identical to his own, than when some or all other agents hold different beliefs.

Proof: Consider wlog agent 1. Fix any belief for agent 1.
Consider first the case that all agents have beliefs that are identical to that of agent 1. Let \( \hat{x} = (\hat{x}_1, ..., \hat{x}_K) \) be the set of actions that maximize \( E_i[\Pi] \) under that belief. I claim that \( \hat{x} = (\hat{x}_1, ..., \hat{x}_K) \) is a Pareto-efficient Nash equilibrium. That it is a Nash equilibrium follows from the fact that for \( \hat{x} - i \) given, \( \hat{x}_i \) maximizes \( E_i[\Pi] \) and thus \( \alpha_i E_i[\Pi] \). Furthermore, if this were not a Pareto-efficient equilibrium, then there existed some equilibrium that gives everyone a weakly higher payoff and at least one agent a strictly higher payoff. This would correspond to a set of actions that give a strictly higher \( E_i[\Pi] \) than \( \hat{x} \), contradicting the definition of \( \hat{x} \). It follows that, under A1, whenever all agents have the same beliefs as agent 1, then the latter’s payoff is \( E_1[\Pi(\hat{x})] \).

But now the proposition follows immediately by the following argument. Fix any set of beliefs for the other agents. Let \( \tilde{x} \) be the actions of any Nash equilibrium that corresponds to this set of beliefs. If \( E_1[\Pi(\tilde{x})] > E_1[\Pi(\hat{x})] \), then that contradicts the definition of \( \hat{x} \) above. This proves the proposition. ■

On the one hand, this result is very strong in that it says that it is always better to associate with people with beliefs that are identical to your own. On the other hand, the result has a key weakness in the fact that it is purely global. It does not say anything about local changes, i.e. whether it is also better to associate with people whose beliefs are a bit more similar to yours. In fact, much more specific assumptions are needed to get such local results. Since the latter are what we need for comparative statics, the next sections study this issue.

Before we continue, however, note that the above formulation and result can be readily adapted to conclude that it is optimal to work with others with values (or utility functions) that are identical to your own. The alternative model in subsection 2.4.3, however, is more fit to this purpose.

### 2.2 Local Results

While the global result is suggestive and encouraging, it does not give us much leverage in terms of analysis. To that purpose, a local result is necessary, in the sense of ‘agent \( i \)’s payoff increases when agent \( j \)’s belief moves closer to \( i \)’s’. Such result does not hold in the same generality, as can be seen from the following example. Let in the setting of subsection 2.1 \( \Pi(x_1, x_2, x) = \pi_1(x_1, x) + \pi_2(x_2, x) \).

The \( \pi_i \) are identical and represented by the upper part of figure 1. As this figure indicates, the payoff contribution \( \pi_i \) of agent \( i \) is maximal when \( \hat{x}_i = x \). More important for this example, however, is the fact that his contribution is higher when \( \hat{x}_i = x + \delta \) than when e.g. \( \hat{x}_i = x + \delta/2 \). The lower part of figure 1 represents the belief distribution of agent \( i \). Consider now a firm with 2 employees who share equally in \( \Pi \). In that case, employee 1 prefers to have a colleague with belief

---

**Figure 1:** A setting in which agents do not necessarily prefer to associate with someone with similar beliefs. The function in the upper part of the graph represents the payoff contribution of an agent in function of his action and the true state of the world. The function in the lower part is the agent’s belief.
functions.

This means that the result does generalize to specific cases where agents have different utility
let, from the agent’s perspective,
the second degree Taylor expansion of Π around the mean of his own distribution. In particular,
suggests a lot about the possibilities and difficulties for doing so. Assume that each agent knows only
Note, for further reference, that the result would also hold if Π = ∑x → x.

To see now how
argmax
comparative statics) that the optimal
not of any other

I first want to show that
Proof : I first want to show that j’s optimal action increases with xj. Note to that purpose
that Ej[Π] = ∑N j=1 Ej[πj(x, xj)] = ∑N j=1 ∫ πj(x, xj)f(x − xj) dx. Since all the other terms are
independent of xj, maximizing Ej[Π] is equivalent to maximizing ∫ πj(x, xj)f(x − xj) dx, which
be written ∫ πj(u + xj, xj)f(u) du. It follows that the optimal xj is a function of xj, but
not of any other xj. Moreover, when πj is supermodular, it follows immediately (by monotone
comparative statics) that the optimal xj increases in xj. In the other case of A2, let
argmaxxj ∫ πj(u − xj)f(u) du, then xj(xj) = xj,0 + xj, so that xj(xj) is again increasing in xj.

To see now how Ei[Π(x)] changes with changes in xj, it is sufficient to look at the term ∫ πj(xj, xj, x) f(x − xj) dx since xj is function of xj and xj does not depend on any actions xj of the other agents. Note that xj(xj) = argmaxz ∫ πj(x, z) f(x − xj) dx. Moreover, by strict quasi-concavity, if z < xj(xj), then ∫ πj(x, z) f(x − xj) dx increases when z → xj(xj). But this means that, when
xj(xj) < xj(xj) (or xj < xj) then ∫ πj(xj, xj, x) f(x − xj) dx increases when xj(xj) → xj(xj) (or
xj(xj) → xj(xj)). This proves the first part of the proposition. The second is completely analogous.

Note, for further reference, that the result would also hold if Π = ∑N j=1 aijπj(x, xj) with aij ≥ 0. This
means that the result does generalize to specific cases where agents have different utility
functions.

More general functions Could this proposition be further generalized? The following result
suggests a lot about the possibilities and difficulties for doing so. Assume that each agent knows only
the second degree Taylor expansion of Π around the mean of his own distribution. In particular,
let, from the agent’s perspective,

Π = ∑N i=1 ai(xj − x) + ∑N i=1 N ai ai(xj − x)2 + ∑N i=1 N i=1 N ai (xj − xj)(xj − xj)

6
Proposition 3 Assume that \( \Pi \) is concave in the decision variables.

- Let there be two agents \((N = 2)\). Profit \( \Pi \) increases as \( \bar{x}_2 \downarrow \bar{x}_1 \) when \( \bar{x}_2 > \bar{x}_1 \) or \( \bar{x}_2 \uparrow \bar{x}_1 \) when \( \bar{x}_2 < \bar{x}_1 \).

- Let there be 3 agents \((N = 3)\). There exists a set of parameters \( \mathbf{a} \) such that \( \Pi \) sometimes increases and sometimes decreases, depending on the value of \( x_3 \), as \( \bar{x}_2 \downarrow \bar{x}_1 \) when \( \bar{x}_2 > \bar{x}_1 \) or \( \bar{x}_2 \uparrow \bar{x}_1 \) in the other case.

- Let there be 3 agents \((N = 3)\). Let \( x_3 \) be drawn from a distribution that is symmetric around \( \bar{x}_1 \). In expectation (over \( x_3 \)), profit \( \Pi \) increases as \( \bar{x}_2 \downarrow \bar{x}_1 \) when \( \bar{x}_2 > \bar{x}_1 \) or \( \bar{x}_2 \uparrow \bar{x}_1 \) when \( \bar{x}_2 < \bar{x}_1 \).

Proof: Note first that for any distribution \( f(x) \),

\[
E[\Pi] = \sum_{i=1}^{N} a_i(x_i - \bar{x}) + \sum_{i=1}^{N} \frac{a_{ii}}{2}(x_i - \bar{x})^2 + \sum_{i=1}^{N} \sum_{j=1}^{N-1} a_{ij}(x_i - x_j)(x_i - x_j) + \sum_{i=1}^{N} \sum_{j=1}^{N-1} \frac{a_{ij}}{2} \sigma^2
\]

with \( \bar{x} \) and \( \sigma^2 \) respectively the mean and the variance of \( f \). Since the last term is constant, it suffices to consider all but the last term in this analysis.

Consider the case with \( N = 2 \) and normalize \( \bar{x}_1 = 0 \). Tedious calculus and algebra shows that

\[
\frac{dE_f(\Pi)}{dx_2} = 2\bar{x}_2 a_{11} \frac{a_{22}^2 + 2a_{12} a_{22} + a_{12}^2}{a_{22} a_{11} - a_{12}^2} = 2\bar{x}_2 a_{11} \frac{(a_{22} + a_{12})^2}{a_{22} a_{11} - a_{12}^2}
\]

which has the opposite sign of \( \bar{x}_2 \), where I use the fact that \( a_{22} a_{11} - a_{12}^2 \geq 0 \) and \( a_{ii} < 0 \) since \( \Pi \) is concave in the actions. It follows that profits increase when \( \bar{x}_2 < \bar{x}_1 = 0 \) and \( \bar{x}_2 \uparrow \bar{x}_1 \), and analogous when \( \bar{x}_2 \downarrow \bar{x}_1 \) and \( \bar{x}_2 > \bar{x}_1 \). This proves the first part of the proposition.

Consider next the case with \( N = 3 \) and normalize again \( \bar{x}_1 = 0 \). In this case

\[
\frac{dE_f(\Pi)}{dx_2} = \frac{A\bar{x}_2 + B\bar{x}_3}{\Delta}
\]

with \( \Delta \) the determinant of the quadratic form (which is negative when the function is concave) and \( A \) and \( B \) independent of \( \bar{x}_2 \) and \( \bar{x}_3 \). Furthermore,

\[
A = (a_{11} a_{33} - a_{13}^2)(a_{22}^2 + a_{12}^2 + a_{23}^2 + 2a_{22} a_{23} + 2a_{22} a_{12} + 2a_{12} a_{23}) = (a_{11} a_{33} - a_{13}^2)(a_{12} + a_{22} + a_{23})^2 > 0
\]

where I use again the fact that \( (a_{11} a_{33} - a_{13}^2) > 0 \). When e.g. \( a_{ii} = 1 \), \( a_{23} = 0 \) and \( a_{12} = a_{13} = 1/2 \), then \( B/\Delta = 9/4 \neq 0 \). The latter implies (e.g. at \( \bar{x}_2 = \bar{x}_1 = 0 \)) that profits will sometimes increase and sometimes decrease as \( \bar{x}_2 \) moves away from \( \bar{x}_1 \), depending on the value of \( \bar{x}_3 \). This confirms the second part of the proposition. Finally, since the derivative is linear in \( \bar{x}_3 \), if \( \bar{x}_3 \) is symmetrically distributed around zero, then the last term will drop in expectation, and we get again that profits increase when \( \bar{x}_2 \) moves closer to \( \bar{x}_1 \). The key reason why the result fails in the second part of the proposition is the interaction between the actions of the agents. A change in \( \bar{x}_i \) changes the optimal actions of all agents, through the equilibrium interaction between \( x_i \) and the other agents’ actions. Sometimes the impact on profits of a change in \( x_j \) may be much larger than the impact of the direct change in \( x_i \).
Consider also the following example. Let there be 3 agents, and let the profit function be

\[ \Pi = A - \frac{(x_2 - x)^2}{2} - \frac{(x_3 - x)^2}{2} - \frac{(x_2 - x_3)^2}{2} \]

Note that agent 1’s action does not play any role. The optimal actions for agents 2 and 3 are \( \hat{x}_2 = \frac{2x_1 + x_3}{3} \) and \( \hat{x}_3 = \frac{x_1 + 2x_3}{3} \). Take now the position of agent 1 with \( x_1 = 0 \). The agent’s payoff is then

\[ A + \frac{1}{3} \left[ -\frac{x_2^2}{3} - \frac{x_3^2}{3} + \frac{x_2x_3}{3} \right] \]

so that

\[ \frac{\partial \Pi}{\partial x_2} = \frac{1}{3} \left[ -2x_2 + x_3 \right] \]

It follows that when \( 0 < x_2 < \frac{x_3}{2} \), then profits increase as \( x_2 \) increases, i.e. as \( x_2 \) moves further away from \( x_1 \). In this case, the gain in profits from moving \( x_2 \) closer is more than compensated by the loss caused by increasing the distance between \( x_2 \) and \( x_3 \).

### 2.3 Effort and heterogeneity

As mentioned earlier, the level of generality of the above conclusions may be somewhat surprising. It is easy to imagine situations where it seems intuitively optimal to work with people with different beliefs. This section shows that that is indeed the case when incontractible effort is important. In particular, I identify two situations where it might be optimal for an agent to associate with someone with different beliefs:

1. When agents can spend private effort to obtain new information and that information cannot be manipulated.

2. When it is optimal to experiment and agents can spend private effort that is complementary to the probability of success.

The first case is analyzed in detail in Van den Steen (2002b). That paper shows that disagreement, in the sense of differing priors, gives agents extra incentives to collect information. The reason is that each agent is convinced that the new information will confirm his point of view and convince the other agents.

For the second case, consider the following setting. A firm consists of 2 agents, denoted 1 and 2, who each can spend effort on a project. There are two project types, \( A \) and \( B \), only one of which will be successful. Projects only generate revenue when successful, with revenue depending on the effort of the agents. In particular, project \( X \) generates a revenue \( \max_{i \in \{1,2\}} e_{i,X} \), where \( e_{i,X} \) denotes the effort of agent \( i \) on project \( X \). The cost of effort to an agent is \( \frac{(e_{i,A} + e_{i,B})^2}{2} \). The agents choose their effort simultaneously, but coordinate on a Pareto-efficient equilibrium when one exists. The probability that each project is successful is unknown, but agents have subjective beliefs in the sense of differing priors. Agent \( i \) believes that project \( X \) is successful with probability \( p_{i,X} \). We have of course that \( p_{i,A} = 1 - p_{i,B} \). It is straightforward that there always exists a Pareto-optimal equilibrium. In that equilibrium the agents spend effort on different projects. In particular, if \( i = \arg\max_{i \in \{1,2\}} \left[ \max_{X \in \{A,B\}} p_{i,X} \right] \), then agent \( i \) spends effort on the project he believes in most,
while the other agent works on the other project. In this case, each agent prefers to work with someone who has completely opposite beliefs from his own. The reason is that such colleague will spend lots of effort on that project since he believes it is the right one. Since the cost of effort is private, this can only benefit the focal agent.

2.4 Remarks

Before moving to the organizational implications of this theory, some extra remarks are in order.

2.4.1 Efficiency

An important consideration is, of course, whether this tendency to associate with similar others is efficient. The answer depends on the notion of efficiency used (Van den Steen 2002c). With differing priors, there indeed two ways to define efficiency. The first notion, ‘subjective’ efficiency, evaluates utilities from the subjective perspective of the agent. An agent is better off when he thinks he is, using his own subjective beliefs. The notion of efficiency, ‘reference’ efficiency, measures utility using some kind of reference belief. This can, for example, be the belief of the social planner or that of the researcher. In the case of an additively separable payoff function that is common to all agents, closer beliefs may be subjectively efficient but are never reference efficient.

Both these notions of efficiency have their relevance and applications. Subjective efficiency, for example, is relevant to see whether agents could improve their situation when improving e.g. contractibility. Reference efficiency is what a social planner or society at large are typically interested in.

2.4.2 Agreement and disagreement in groups

One way of describing the analysis in this section is that it studies the utility externalities of differences in beliefs in joint undertakings. Since utility is what drives people, this analysis also has implications regarding disagreement on actions and the ensuing influence activities. While these implications might seem a bit trivial in nature, they are potentially very useful from an empirical perspective, since they might provide indirect measures of belief differences.

2.4.3 Alternative formulations

The models in this section all assumed that each agent undertakes an action independently of all other agents. Very often, however, agents have to jointly undertake one and only one action. In such case, the decision making process determines how the different beliefs get integrated to a joint action. This perspective is probably the most fruitful to discuss culture in terms of differing values or utility functions.

3 Definitions of culture and cultural strength

While there is considerable variation in the literature, it is fair to say that by far the leading definition of corporate culture is that of culture as shared assumptions, beliefs, or values⁴. In a

⁴Although a complete review is outside the scope of this work, it seems that at least 70 to 80% of the management and organizational literature subscribes to the view of culture as shared beliefs and values. The other 20 to 30%, however, is a complete amalgam or avoids any definition at all.
contribution that preceded the main culture literature, Burns and Stalker (1961) define culture in the context of organic organizations as ‘a dependable constant system of shared beliefs’. Peters and Waterman (1982) define it as ‘shared values’ but elsewhere explain that this includes ‘basic beliefs’. Donaldson and Lorsch (1983) is often considered a seminal work on corporate culture, although it talks instead about managerial beliefs.

Probably the most cited perspective on corporate culture is that of Schein (1985). He defines culture as having three levels. The most visible, but most superficial, level is that of culture as a pattern of behavior. It is ‘the way things are done around here’, the norms, the stories. The behavioral patterns are driven by the second level of culture, which are the shared values. Shared values are, on their turn, driven by the third and most fundamental level of culture: shared assumptions. Note that beliefs and values are, both in theory and in practice, very closely related. Most of the management literature tends to use the terms interchangeably. Kotter and Heskett (1992), for example, base their definition on Schein (1985), but eliminate the distinction between the second and the third level.

Given the definition of corporate culture as shared beliefs, the literature gives essentially two interpretations to the notion of ‘strength’ of such culture: how similar it is internally and how different it is externally, i.e. how different these common beliefs are from the beliefs of the population at large. I will reserve the term ‘strength’ for the first, i.e. for the degree of homogeneity of beliefs within the firm, and use ‘distinctiveness’ for the degree of difference with the population at large. With $d$ a (semi-)metric on the real line, I will define the strength of corporate culture, $\tau$, as the inverse of the average distance between two randomly selected member of the organization.

$$\frac{1}{\tau} = \frac{1}{I(I - 1)} \sum_{i=1}^{I} \sum_{j=1}^{I} d(\pi_i, \pi_j)$$

Of course other measures, such as the range, are possible.

4 Formation of Culture

Employees’ tendency to associate with similar others will lead to a certain level of homogeneity of beliefs or values within a firm. As argued above, this corresponds to the notion of ‘corporate culture’. The next sections study the formation and implications of such culture.

4.1 Methodological issues

There are, however, two important things to consider before starting this analysis.

The first issue is the question how to model the hiring process. One could, for example, imagine any of the following processes:

---

5Schein’s formal definition is as follows: ‘A pattern of shared basic assumptions that the group learned as it solved problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems.’ While this definition might seem to contradict the current paper’s focus on recruitment, Schein at a later stage in that book states that ‘One of the most subtle yet most potent ways through which cultural assumptions get embedded and perpetuated is the process of selecting new members. ... This cultural embedding mechanism is subtle because it operates unconsciously in most organizations. Founder and leaders tend to find attractive those candidates who resemble present members in style, assumptions, values, and beliefs.’

6As before the analysis will be done in terms of differing beliefs, but can easily be adapted to heterogenous values or utility functions.
1. Someone starts the firm and that person, the ‘founder’, does all the hiring and firing. This would be typical for an early-stage startup.

2. The founder (or other person in charge) designates a current member (or group of members) of the firm to do the hiring. This is typical for a larger firm.

3. All the current members of the firm decide jointly on hiring. In this case, the decision could be taken, for example, by unanimity or by a majority vote. Small professional firms and academic departments often function this way.

4. A random selection of current members decides on hiring, again by unanimity or by majority vote. Large professional firms typically work this way.

Sometimes it is difficult for the firm to figure out the beliefs of potential new employees but the latter have a quite clear idea of the beliefs of the firm’s current employees. In that case, the firm cannot select on beliefs and therefore hires whoever comes its way (as long as adding an employee increases profits). In that case, it is up to the potential new employees to decide whether to join or not. This is a completely different ‘hiring’ process than the ones above.

Any of these hiring processes will typically have different implications for the company’s culture. Figuring out these differences is a research program on itself, far beyond the scope of the current paper. It poses, though, the problem to select a hiring process that is interesting to look at and analytically tractable. While more work will be done in future versions of this paper, the current focus is on the case where the founder does all the hiring.

A second issue is the dynamic nature of the process, which can make the analysis enormously complex. Consider, for example, a situation in which the joining is decided by complete unanimity i.e. all current plus the new employee must approve the new employee’s joining. At each stage an agent’s decision must also consider how this new member will vote in the future. Such considerations quickly get extremely involved. Upon hiring the second employee, for example, the first employee has to consider how that second employee will vote when the fourth employee comes up. But this requires the first employee to consider how the second thinks about influence that fourth employee will have when the fifth employee comes up. It is clear that a completely general analysis is nearly impossible. It is also not clear whether this would be a very accurate analysis of the true hiring process.

This paper takes a number of different approaches on this dynamic issue.

The next section simplifies the problem in two ways. For one thing, it only considers the hiring of one extra agent, i.e. it studies only one step of the full dynamic process. For another, it assumes that this is the last agent or that $\delta \to 0$. Dynamic considerations will therefore not play any role. As a consequence, however, the analysis will not have to say much about the overall pattern of beliefs or about the cultural strength. Nevertheless, it is very useful, as I discuss later.

Section 6 considers the full dynamic process in the limit as $\delta \to 1$. This is a good approximation of a situation with very rapid hiring. More importantly, it also describes a situation in which all agents reconsider their hiring/joining decisions each time a new employee might get hired. Employee by employee, the manager decides whether to keep (‘re-hire’) the employee while the employee decides whether to stay (‘re-join’), both knowing exactly what the beliefs of all other employees are.

---

7 Note that even with $\delta \to 0$, there are dynamic issues in a general analysis. In particular, the hiring of an agent at time $t$ changes the configuration, and thus the hiring/joining decisions, at time $t + s$. It follows that the sequence in which agents get hired still is a consideration.
Section 7 does a brute-force analysis of the complete dynamic problem for a limited number of agents and an extremely simple specification. It does nevertheless give very interesting insights on the dynamics of hiring.

A useful approach that is not followed in this paper is to look at ‘stable associations’. In particular, say that \( I \subset K \) is a stable association if none of the current members would be better off leaving the organization and none of the current non-members would be better off joining. The reason for not taking this approach here (for now) is that there often is a multiplicity of stable associations for a given set of beliefs. Unless we introduce some selection rule to obtain uniqueness, this makes it very difficult to talk about the average properties of the configurations, which is what we need for empirical work. On the other hand, it seems that many selection rules reduce this problem to the dynamic case with \( \delta \to 1 \). For example, if I were to use the rule ‘select the configuration that agent \( i \) prefers’, then that is identical to a hiring process with \( \delta = 1 \) where \( i \) is the founder and makes all hiring decisions. This is exactly the approach followed in section 6.

Apart from these dynamic issues, there is also the question how to specify the agents’ utility functions. Section 2.2 made already clear that we can’t get by with a fully general payoff function. To get insightful comparative statics, however, I will impose still more structure. The general model for the coming sections is discussed below.

### 4.2 Setup and parametrization

Consider \( K \) agents who could join a firm. As before, each agent \( i \) has to undertake an action \( x_i \in \mathbb{R} \). The firm’s total profit is the sum of the contributions of the participating agents. Agent \( i \)'s contribution is \( \pi(I) + \gamma g(x - x_i) \), with \( x \) denoting the unknown state of the world about which agents will have subjective beliefs and \( I \leq K \) the number of members of the firm. The total firm profit is thus

\[
\Pi = \sum_{i=1}^{I} \pi(I) + \gamma g(x - x_i)
\]

Note that this profit function is additively separable in the agents’ decisions. The parameter \( \gamma \) measures how much the agents’ decisions affect the profit, i.e. \( \gamma \) measures the importance of discretion. Sometimes I will assume that there are scale effects, i.e. that \( \pi(I) \) increases in \( I \).

Instead of assuming that each agent gets a fixed share of the total profits, I assume here agent \( i \)'s payoff is a convex combination of his own contribution and his equal share of the joint profit. So agent \( i \)'s utility is

\[
u_i = \theta [\pi(I) + \gamma g(x - x_i)] + (1 - \theta) \left[ \frac{1}{I} \sum_{j=1}^{I} \pi(I) + \gamma g(x - x_j) \right] \\
= \pi(I) + \theta \gamma g(x - x_i) + (1 - \theta) \gamma \frac{1}{I} \sum_{j=1}^{I} g(x - x_j)
\]

 Appropriately adapted versions of propositions 1 and 2 still apply in this case. The reason for using this slightly more general specification is that it allows me to derive comparative statics on the parameter \( \theta \), which measures how important individual performance is for the agent’s payoff, relative to the importance of group performance. The reason for introducing this only now is that it would have complicated the analysis in section 2 without adding any insights.
As mentioned earlier, the true state of the world $x$ is unknown to the agents, but they have subjective beliefs. In particular, each agent $i$ has a subjective belief $F_i$ with density $f_i$. Assume furthermore $f_i(x) = f_j(x + x_j - x_i) = f(x - x_i)$, so that the agents’ beliefs are identical up to their mean. In earlier versions of this paper, both the $g$-functions and the densities were allowed to vary by agent, but that doesn’t add much beyond a lot of indices.

Note that $E_i[g(x - x_j)] = \int g(x - x_j) f_i(x - x_i) dx = \int g(x - (x_j - x_i)) f(x) dx$. Define $D(z) = \int g(x - z) f(x) dx$ and make the following parametric assumption

**Assumption 3** Let $D$ be continuous and strictly quasiconcave.

This assumption will be satisfied when $g$ is continuous and both $g$ and $f$ are strictly quasi-concave. It also implies that $D$ has a unique maximizer, denoted $\eta$. This implies that $\hat{x}_i = x_i + \eta$ and that at the optimum agent $i$’s expected utility is

$$E_i[u_i] = \pi(I) + \gamma \left[ \theta \gamma D(\eta) + (1 - \theta) \frac{1}{I} \sum_{j=1}^{I} D(x_j - x_i + \eta) \right]$$

or, using $\hat{D} = D(\eta)$ and $d(z) = D(\eta) - D(z + \eta)$,

$$E_i[u_i] = \pi(I) + \gamma \hat{D} - \gamma (1 - \theta) \frac{1}{I} \sum_{j=1}^{I} d(x_j - x_i)$$

Note that if $D$ is symmetric then so will be $d$. In that case, $d$ is a semi-metric function\(^8\). Appendix B discusses the possible origins of such distance payoff functions.

### 5 Static Preferences for Association

Consider now a situation where the firm currently consists of agents \( \{1, 2, \ldots, I\} \), and there is one more potential employee, denoted \((I + 1)\). In this case, the agents do not have to consider any future consequences of hiring this employee, resp. joining the firm. Such static hiring/joining preferences are important for three types of analyses:

1. The literal case that there is only one more potential employee.
2. A dynamic hiring process with discount rate $\delta \to 0$.
3. The question of stability of a specific configuration of beliefs, as discussed above in section 4.

Note indeed that in the latter case, there are also no dynamic considerations in determining the gain from hiring/joining.

I now first consider how the current members of the organization gain or lose from hiring agent \((I + 1)\), and then take the perspective of agent \((I + 1)\) on him joining the firm.

---

\(^8\)A semi-metric function satisfies all properties of a metric, except for the triangle inequality.
5.0.1 Hiring preferences

Consider thus first how ‘insiders’ look at ‘outsiders’, i.e. how the current members of the firm feel about (and thus would decide on) a specific new member joining. Let \(1 \leq i \leq I\), so that agent \(i\) is currently a member of the organization.

**Proposition 4a** Agent \(i\) wants \((I + 1)\) to join if

\[
d(\pi_{(I+1)} - \pi_i) \leq (I + 1) \left[ \frac{\pi((I + 1)) - \pi(I)}{\gamma(1 - \theta)} \right] + \frac{1}{I} \sum_{j=1}^{I} d(\pi_j - \pi_i)
\]

**Proof:** Agent \(i\) want \((I + 1)\) to join if

\[
\pi(I) + \gamma \hat{D} - \gamma(1 - \theta) \frac{I}{I} \sum_{j=1}^{I} d(\pi_j - \pi_i) \leq \pi(I + 1) + \gamma \hat{D} - \gamma(1 - \theta) \frac{I+1}{I+1} \sum_{j=1}^{I+1} d(\pi_j - \pi_i)
\]

or

\[
(I + 1)\pi(I) - \gamma(1 - \theta) \frac{I+1}{I} \sum_{j=1}^{I} d(\pi_j - \pi_i)
\]

\[
\leq (I + 1)\pi(I + 1) - \gamma(1 - \theta) \sum_{j=1}^{I} d(\pi_j - \pi_i) - \gamma(1 - \theta)d(\pi_{(I+1)} - \pi_i)
\]

or

\[
d(\pi_{(I+1)} - \pi_i) \leq (I + 1) \left[ \frac{\pi((I + 1)) - \pi(I)}{\gamma(1 - \theta)} \right] + \frac{1}{I} \sum_{j=1}^{I} d(\pi_j - \pi_i)
\]

Note that this condition specifies that \(\pi_{(I+1)} \in [\pi_i - \epsilon_i, \pi_i + \epsilon_i]\) with \(\epsilon_i, \pi_i \geq 0\). Call \(\pi_i + \epsilon_i\) agent \(i\)’s ‘hiring radius’ \(R_i\).

**Proposition 4b** An agent’s hiring radius \(R_i\) decreases in \(\gamma\), increases in \(\theta\), decreases in the average semi-distance to the current members of this firm \(\frac{1}{I} \sum_{j=1}^{I} d(\pi_j - \pi_i)\), and increases in the scale effects of \(\pi\), \(\pi(I + 1) - \pi(I)\).

**Proof:** The proof follows directly from monotone comparative statics on \(R_i\).
With respect to the first two comparative statics, note that professional firms, which typically have very strong cultures, also have a lot of shared rewards, while incontractible decisions are key.

The third comparative static is probably the most surprising. It suggests that cultural strength will be path dependent: if the organization starts (by pure luck) with a strong culture in the sense of homogenous beliefs, then it members will have a tendency to be more strict in their hiring. This effect is caused by the fact that the existing members have a high subjective payoff from having such homogenous beliefs, and that they have to share that payoff when they admit a new member. They will only be willing to do so when that member is also sufficiently ‘close’ in terms of beliefs.

5.1 Joining preferences

Consider next under what conditions agent \((I + 1)\) will want to join the firm. Let the outside option of the agent be \(w\).

**Proposition 4c** Agent \((I + 1)\) will want to join if

\[
\frac{1}{(I + 1)} \sum_{j=1}^{I} d(\bar{x}_j - \bar{x}_{(I+1)}) \leq \frac{\pi(I + 1) + \gamma \hat{D} - w}{\gamma(1 - \theta)}
\]

**Proof:** Agent \((I + 1)\) will want to join if

\[
w \leq \pi(I + 1) + \gamma \hat{D} - \gamma(1 - \theta) \frac{1}{(I + 1)} \sum_{j=1}^{I} d(\bar{x}_j - \bar{x}_{(I+1)})
\]

or

\[
\frac{1}{(I + 1)} \sum_{j=1}^{I} d(\bar{x}_j - \bar{x}_{(I+1)}) \leq \frac{\pi(I + 1) + \gamma \hat{D} - w}{\gamma(1 - \theta)}
\]

Note that this describes a sphere around \(\bar{x}_{(I+1)}\) in this semi-metric \(d\). Denote the right hand side to be the ‘joining radius’ of agent \((I + 1)\).

**Proposition 4d** The joining radius of agent \((I + 1)\) decreases in \(\gamma\), increases in \(\theta\), increases in \(\pi(I + 1)\), and decreases in the outside option \(w\).

**Proof:** The proof follows again immediately from monotone comparative statics on the definition of the joining radius.

These results are again intuitive. Take for example the outside wage \(w\). As this wage increases, the agent’s outside option improves, which makes him more picky as to what organizations he wants to join. This corresponds with a decrease in the joining radius.

6 A Dynamic Model for \(\delta \to 1\)

This section studies a dynamic model with a fairly general parametrization, but limits itself to the case that \(\delta \to 1\). Moreover, I will assume that the beliefs of all agents are randomly and
independently drawn from some distribution of beliefs at the start of the game. It follows that all agents know up front all beliefs of the others.

Consider in particular the following game. There are $I$ agents, designated agents 1 through $I$. Agent 1 is the ‘founder’ of the organization with belief $\pi_1 = 0$. Agents 2 through $I$ are potential employees. Their beliefs are independent draws from some distribution $H$ with density $h$. I make moreover the following assumption

**Assumption 4** $H$ is symmetric around $\pi_1 = 0$. The function $d$ is symmetric around 0.

For the second part of the assumption, it is sufficient that the $f$ and $g$ defined in section 4.2 are symmetric around zero.

The game consists of $I + 1$ stages, but there is no discounting (i.e. $\delta = 1$). In stage 1, all beliefs get revealed. Stages 2 through $I$ are identical. In particular, in stage $k$

1. The founder decides whether or not to allow agent $k$ in the firm.

2. If the founder allows agent $k$ in the firm, then agent $k$ has to decide to accept the offer or not.

In stage $I + 1$, the payoffs get realized. In particular, all agents who were not allowed to join or who refused to join, get their outside option $w$. The founder and all employees who accepted his offer to join produce according to the specification in section 4.2. In particular, let $J$ denote the set of employees who are in the firm, then the expected payoff of agent $i \in J$ is

$$ E_i[u_i] = \pi(J) + \gamma \bar{D} - \gamma(1 - \theta) \frac{1}{J} \sum_{j \in J} d(\pi_j - \pi_i) $$

where I abuse notation $J$ to denote also the cardinality of the set $J$. To further simplify the analysis, make the following assumption.

**Assumption 5** The outside wage $w = -\infty$, so that no agent will refuse to join.

An important implication of this assumption is that the order of the employees doesn’t matter. This simplifies the proof a lot. Assume finally that the founder when indifferent admits the potential member.

**A note on interpretation** This game structure can be interpreted as being equivalent to a completely general dynamic game with employment at will. In particular, assume that the founder can hire and fire at will, and that individual agents can refuse to join or decide to leave the firm at will. Assume that at randomly determined moments, new potential employees will appear. The beliefs of these employees are still independent draws from $H$, but they are revealed only when the employee ‘appears’ (and thus not up front). After the agent’s belief is publicly observed, the founder can fire some of the firm’s current members and then sequentially ask some of the non-associated agents (including those who were just fired) to join the firm. Finally, any employee can leave the firm. At the end of the period the firm produces a payoff. Agents never leave the game. So any agent who is never asked to join the firm, who refused to join or who was fired is just a non-associated agent for the rest of the period and produces some outside value $w$. Any equilibrium of this game is equivalent with one in which the founder fires all employees at the time a new employee appears and then re-hires one by one. This is exactly the game described above.
**Results** Let $\bar{J}(I) \subset I$ denote the subset of employees that, when joining the firm, maximizes the payoff from the perspective of the founder. The equilibrium of this game is then simply that the founder makes offers to all members of $\bar{J}(I)$ and all these agents accept.

**Lemma 1** There exists $\bar{d}$ such that $\bar{J}(I) = \{i \in I \mid d(\bar{x}_i, \bar{x}_1) \leq \bar{d}\}$.

The proofs of this and the next lemma are in appendix.

**Lemma 2** The ‘radius’ $\bar{d}$ decreases as $\gamma$ and $(1 - \theta)$ increase

For the next result, remember that we defined the cultural strength of an organization to be the inverse of the average distance between two randomly selected members:

$$\frac{1}{\tau} = \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{j=1}^{I} d(\bar{x}_i, \bar{x}_j)$$

**Proposition 5** The expected inverse cultural strength, $E[1/\tau]$, decreases as $\gamma$ or $(1 - \theta)$ increase.

**Proof**: Fix some founder $\bar{x}_1$. Let $k_J(x)$ for $J \in \sigma(I)$ be the indicator function for the fact that the subset $J$ are exactly the agents that get hired by the founder for this configuration $x$. The expected inverse cultural strength can then be written:

$$E[1/\tau] = \sum_{J \in \sigma(I)} \int_x \left[ \frac{1}{J(J-1)} \sum_{i \in J} \sum_{j \in J} d(\bar{x}_i, \bar{x}_j) \right] k_J(x) f(x) \, dx$$

with $h(x)$ the (symmetric) density of $x$. Let $S_J = \{x \mid k_J(x) = 1\}$. For all $x \in S_J$, renumber for what follows the indices such that $d(\bar{x}_i, \bar{x}_1) \leq d(\bar{x}_{i+1}, \bar{x}_1)$. Note that agent $J$ is the member of the firm whose belief $\bar{x}_J$ differs most from that of the founder. Lemma 2 implies that as $\gamma$ changes, some agents with beliefs $\bar{x}_J$ will drop out of the firm (which may lead to more agents dropping out; I do the proof here under the assumption that only one agent drops out; it is straightforward but tedious to extend it for multiple agents dropping out). We have to show that this dropping out reduces $E[1/\tau]$. Let now $\partial S_J$ be the subset of $S_J$ such that the founder is just indifferent whether or not to hire agent $J$ and $\partial S_{J,a} = \{x \in \partial S_J \mid \bar{x}_J = \bar{x}_1 - a\}$. We can write:

$$E[1/\tau] = \sum_{J \in \sigma(I)} \int_{S_J \setminus \partial S_J} \left[ \frac{1}{J(J-1)} \sum_{i \in J} \sum_{j \in J} d(\bar{x}_i, \bar{x}_j) \right] h(x) \, dx + \sum_{J \in \sigma(I)} \int_{\partial S_J} \left[ \frac{1}{J(J-1)} \sum_{i \in J} \sum_{j \in J} d(\bar{x}_i, \bar{x}_j) \right] h(x) \, dx$$

Fix some $a$ and let $g(x_{-J}) = k_J(x_{-J}, \bar{x}_1 - a) h(x_{-J}, \bar{x}_1 - a)$, which is symmetric around $\bar{x}_1$ in $x_{-J}$.

I am left to show that

$$\int \frac{1}{(J-1)(J-2)} \sum_{i=1}^{J-1} \sum_{j=1}^{J-1} d(\bar{x}_i, \bar{x}_j) \, dG \leq \int \frac{1}{(J-1)J} \sum_{i=1}^{J} \sum_{j=1}^{J} d(\bar{x}_i, \bar{x}_j) \, dG$$
or

\[
\int \frac{1}{(J - 1)(J - 2)} \sum_{i=1}^{J-1} \sum_{j=1}^{J-1} d(\pi_i, \pi_j) \, dG \\
\leq \int \frac{1}{(J - 1)^2} \sum_{i=1}^{J-1} \left( \sum_{j=1}^{J-1} \left( d(\pi_i, \pi_j) + 2d(\pi_i, \pi_1 - a) \right) \right) \, dG
\]

Note that this inequality can be rewritten

\[
\int \sum_{i=1}^{J-1} \sum_{j=1}^{J-1} d(\pi_i, \pi_j) \, dG \leq \int \sum_{i=1}^{J-1} ((J - 2)d(\pi_i, \pi_1 - a)) \, dG
\]

or

\[
\int 2 \sum_{i=1}^{J-1} \sum_{j=1}^{J-1} d(\pi_i, \pi_j) \, dG \leq \int \sum_{i=1}^{J-1} \sum_{j=1}^{J-1} (d(\pi_i, \pi_1 - a) + d(\pi_j, \pi_1 - a))
\]

So that it is sufficient to show for any \( \pi_i \) and \( \pi_j \) that

\[
\int 2d(\pi_i, \pi_j) \, dG \leq \int (d(\pi_i, \pi_1 - a) + d(\pi_j, \pi_1 - a)) \, dG
\]

By the above, it suffices to show that for any two members with (joint) belief distribution \( g(x, y) \):

\[
\int_{\pi_1-a}^{\pi_1+a} \int_{\pi_1-a}^{\pi_1+a} 2d(x, y)g(x, y) \, dx \, dy \leq \int_{\pi_1-a}^{\pi_1+a} \int_{\pi_1-a}^{\pi_1+a} d(x, \pi_1-a) + d(y, \pi_1-a)g(x, y) \, dx \, dy
\]

Notice that by assumption \( g(x, y) \) is symmetric around \( \pi_1 \). It follows that the above left-hand-side can be written

\[
\int_{\pi_1-a}^{\pi_1} \int_{\pi_1-a}^{\pi_1} 2d(x, y) - d(x, \pi_1-a) - d(y, \pi_1-a)g(x, y) \, dx \, dy
\]

or, using symmetry and after some algebra,

\[
\int_{\pi_1-a}^{\pi_1} \int_{\pi_1-a}^{\pi_1} 4d(x - y) - 2d(x - \pi_1 + a) - 2d(y - \pi_1 + a)g(x, y) \, dx \, dy
\]

Note now that, by definition of the integration intervals and of the function \( d \),
\[ d(x - y) \leq d(x - (\overline{x}_1 - a)) \]
\[ d(x - y) \leq d(y - (\overline{x}_1 - a)) \]
\[ |\overline{x}_1 - x| \leq a \text{ so that } d((\overline{x}_1 - x) + (\overline{x}_1 - y)) \leq d((\overline{x}_1 - y) + a). \]
\[ |\overline{x}_1 - y| \leq a \text{ so that } d((\overline{x}_1 - x) + (\overline{x}_1 - y)) \leq d((\overline{x}_1 - x) + a). \]

This implies that the equation is negative. Notice moreover that the inequality will be strict whenever \( g(x, y) \) is continuous and \( a > 0 \).

7  A Complete (but simple) Dynamic Model

This section considers a general dynamic process for an extremely simple parametrization and hiring process. This achieves two things. It confirms, for this particular case, the main comparative statics of section 5. It also gives some new insights in how the dynamic considerations may influence an organization’s culture.

7.1 Setup and parametrization

Consider a situation with 3 agents, denoted 1, 2, and 3. Agent 1 is the founder of the firm and will decide on himself on all the hiring. Agents 2 and 3 are potential employees. Let \( g(x) = A - |x| \) and assume that the agents have degenerate beliefs in the sense that agent \( i \) thinks \( x = \overline{x}_i \) with probability one. In that case,

\[ D(z) = \int A - g(u - z)dF(u) = A - \int |u - z|dF_i(u) = A - |z| \]

This implies that \( \eta = 0, \hat{x}_i = \overline{x}_i, \hat{D} = A, d(z) = A - A + |z| = |z| \). Assume, finally, that \( \pi(I) = \epsilon \left(1 - \frac{1}{T}\right) \) for some \( \epsilon > 0 \). It follows that

\[ E_i[u_i] = \epsilon \left(1 - \frac{1}{T}\right) + \gamma A - \gamma(1 - \theta)\frac{1}{T} \sum_{j=1}^{T} |\overline{x}_j - \overline{x}_i| \]

or, with \( X = \frac{\epsilon}{\gamma(1-\theta)} \)

\[ E_i[u_i] = \left[X \left(1 - \frac{1}{T}\right) + \frac{A}{(1-\theta)} - \frac{1}{T} \sum_{j=1}^{T} |\overline{x}_j - \overline{x}_i| \right] \gamma(1 - \theta) \]

Assume further that \( \overline{x}_2 \) and \( \overline{x}_3 \) are random draws from a uniform distribution on \([\overline{x}_1 - .5, \overline{x}_1 + .5]\), and that \( X < .5 \). Assume finally that an agent’s outside option \( w \) is so small that the agent always wants to join if the founder asks him to (e.g. \( w = -\infty \)).

Assume the following timeline. There are 2 periods. In period 1, agent 2 gets drawn from the uniform distribution. The founder can then offer agent 2 a position in his firm. If he doesn’t do so, then he produces alone that period, agent 2 gets \( w \) for the period, and the period is over. If he does offer 2 a job and 2 rejects, the outcome is the same as above. In either case, agent 2 disappears
from the game at the end of period 1. If he offers 2 a job and 2 accepts, then 2 becomes part of the firm, they produce together and split the profits according to the utility functions described above. In period 2, the firm starts with the same composition as it ended period 1. Agent 3 gets drawn. Period 2 then goes exactly as period 1, only now with agent 3 in 2’s position. The discount factor is $\delta$.

**Results** In this case, the founder will hire agent 2 if $\bar{x}_2 \in [-\tilde{x}_2, \tilde{x}_2]$ with

$$\tilde{x}_2 = (3\alpha - 1) - \sqrt{(3\alpha - 1)^2 - 6\alpha X + 5X^2}$$

where $\alpha = \frac{1+\delta}{\delta}$. Furthermore, if he did not hire agent 2 in the first period, then he hires agent 3 if $\bar{x}_3 \in [-X, X]$, otherwise he hires agent 3 if $-\bar{x}_3 \in \left[\frac{X+\bar{x}_2}{2}, \frac{X-\bar{x}_2}{2}\right]$.

The following proposition confirms essentially the results of the last section

**Proposition 6** The expected inverse cultural strength, $E[1/\tau]$, increases in $X = \frac{\epsilon \gamma (1-\theta)}{\theta}$.

**Proof:** The proof is extremely long and therefore omitted (but can be obtained from the author).

This essentially confirms the results of the last two sections in this dynamic context.

**Path-dependence** An important insight from this dynamic analysis is how both the content and the strength of culture are path-dependent.

The path-dependence of cultural strength had been already suggested in section 5. Note here that, after hiring agent 2, the hiring interval for agent 3 becomes $\bar{x}_3 \in \left[\frac{X+\bar{x}_2}{2}, -\frac{X+\bar{x}_2}{2}\right]$. It follows that the founder will be more picky about which agent to hire when $\bar{x}_2$ happened to be more similar to the founder’s belief. This also suggests that culturally strong firms will be more selective and therefore smaller for a given number of potential hires.

The path dependence of the culture content seems obvious since the founder will hire in an interval around his own belief. It follows that the employees’ beliefs will be clustered around what happened to be the founder’s belief. This confirms in a strong way the notion that an organization’s culture is determined by its founder (Schein 1985).

There are, however, more subtle ways how this path dependence of cultural content can operate. This is especially the case when $w$ is such that potential employees might sometimes reject a job offer from the founder. This will immediately cause again path-dependence in that only employees with beliefs sufficiently similar to these of the organization will join. Note here that all beliefs will count. So even if the hiring is done by completely by the founder, the final pattern of beliefs will also depend on the beliefs of the early hires. The most subtle way is the following conjecture. Assume that, by coincidence, the first three hires are to the left of the founder’s belief. Assuming that there are still many more potential hires out there, the founder will now be biased in favor of potential hires that are also on his left side. The reason is the following: if he were to hire an employee on his right side, that would spread out the beliefs of the organization more than hiring an employee on his left side would. Such spread makes the organization on average less attractive to future new hires.
8 Benefits and Costs of Strong Cultures

The above analysis raises the question when a strong culture is desirable and when it might actually be better to have a diversity of beliefs and values. While providing a complete answer is beyond the scope of this paper, it is probably useful to summarize some results in this direction.

Incentives for information collection, communication, and other effort  Van den Steen (2002b) shows that diversity in beliefs increases the agents’ incentives to collect information. The reason is that each agent is convinced that new data will prove him right and allow him to convince the other agents. On the other hand, disagreement distorts the incentives to communicate. In particular, in trying to convince their colleagues, agents sometimes withhold information that is beneficial but damages their position. Finally, disagreement can reduce the incentives of agents to spend other effort, in particular when the agents have to agree on a course of action and effort is complementary to that action being successful.

Coordination  The coordinating benefits of corporate culture have been widely touted (e.g. Kotter and Heskett 1992). A simple model in Van den Steen (2002a) shows how miscoordination indeed decreases when the employees of the firm have more similar beliefs. Surprisingly, however, this does not necessarily imply that firms will develop a stronger culture when coordination is important. The reason is that as coordination becomes more important, the agents adapt their actions accordingly. This reduces the coordination gain from stronger culture, and thus the incentives to create a strong culture.

Dealing with change and innovation  At first sight, it seems intuitive that culture as homogenous beliefs would impede change and innovation. Some companies with very strong cultures, however, are also very innovative. Apple and 3M are two cases that spring to mind. Peters and Waterman (1982) even argued that a strong culture is the foundation for innovation. The relationship between culture and innovation therefore seems to depend on the content of the beliefs that underly the culture. This is definitely a topic that requires more study.

9 Managing culture

Although the paper has not detailed all relevant aspects, the analysis suggests some ways how companies can affect culture.

Hiring processes  While the specifics have to be researched in more depth, the form of the hiring process will definitely affect the strength of the organizational culture. The hiring processes of professional firms are rather revealing in that respect. Such firms very often have a dual hiring process. Administrative staff, whose actions are fairly routine and do not influence so much the performance or reputation of the firm, are typically hired by a human resources manager. That human resources manager, however, has little, if any, role in the hiring of professionals. Hiring of professionals is nearly always done by other professionals. Often the hiring process is very extensive. Senior people are involved in the hiring of even rookie consultants. The role of such senior people in the process is essentially to see whether the candidate will ‘fit’ the firm. Professional firms with a strong culture also typically spend quite a lot of time explaining their values to new hires. One effect of this communication is to quickly weed out those who aren’t at ease with that value system.
Communicating values and beliefs to candidates and new members The communication of values and beliefs starts actually much earlier than the hiring process. Many companies with strong cultures are quite open about their culture and make it a key part of any presentation about the firm or company. Such communication can again serve as a (self)screening device prior to the real hiring process.

Compensation practices Compensation can also play a role in creating or maintaining the firm’s culture. In this case, low wages can serve the screening function: only employees with the right beliefs will be willing to come to the firm at such low wages. Paying in stock options can help to select people who share your beliefs regarding the right course of action in this industry.

New leadership with strong beliefs The model of vision in Van den Steen (2001) also shows how the beliefs of the current CEO will typically make the firm attractive to people who share that belief and unattractive to others. Hiring a CEO with very strong beliefs can therefore be an effective mechanism to create or change the culture of the company. This has been described in much detail in the management and organizations literature (Schein 1985, Kotter and Heskett 1992). Note that such ‘change management’ will cause turnover since the people who do not agree with the new CEO will consider the firm much less attractive. But if changing the culture is what you want to do, then such turnover might be exactly what you want.

10 Conclusion

This paper showed that organizational culture might be generated as a by-product of the subjectively rational tendency of agents to ally with similar others. It analyzed the consequences of such theory in terms of the determinants of cultural strength and the dynamics of culture formation. The paper concluded by surveying some recent results on the benefits and costs of homogenous beliefs.

Since culture is so important in organization theory, there are many ways in which this theory can lead to further research. The paper hinted at some, such as the influence of alternative hiring processes and the interaction with change and innovation. The reader probably has little difficulty coming up with his own favorite issues.
A Proof of propositions

Note that I assumed $D$ and thus $d$ to be symmetric around 0.

**Lemma 1** There exists $d$ such that $\tilde{J}(I) = \{ i \in I \mid d(\bar{x}_i, \bar{x}_1) \leq \tilde{d} \}$.

**Proof:** Assume (by contradiction) that $i \in \tilde{J}(I)$, $k \notin \tilde{J}(I)$ while $d(\bar{x}_1, \bar{x}_k) < d(\bar{x}_1, \bar{x}_i)$. Given that 1’s payoff function equals

$$E_1[u_1] = \pi(J) + \gamma \hat{D} - \gamma(1 - \theta) \frac{1}{J} \sum_{j=1}^{J} d(\bar{x}_j - \bar{x}_1)$$

the founder could improve his payoff by hiring $k$ instead of $i$. But this implies the lemma. ■

**Lemma 2** The radius $\tilde{d}$ decreases as $\gamma$ or $(1 - \theta)$ increase

**Proof:** Consider the case of $\gamma$. Order (wlog) the agents by their distance from $\bar{x}_1$. For a given $\gamma$, there exists some $n$ such that $i \in \tilde{J}(I) \text{ iff } i \leq n$, by lemma 1.

The change in profit for the founder when the firm goes from $J$ to $(J + 1)$ agents is

$$\left[ \pi(J + 1) + \gamma \hat{D} - \gamma(1 - \theta) \frac{1}{J + 1} \sum_{j=1}^{J+1} d(\bar{x}_j - \bar{x}_1) \right] - \left[ \pi(J) + \gamma \hat{D} - \gamma(1 - \theta) \frac{1}{J} \sum_{j=1}^{J} d(\bar{x}_j - \bar{x}_1) \right]$$

or

$$\left( \pi(J + 1) - \pi(J) \right) + \gamma(1 - \theta) \left[ \frac{1}{J} \sum_{j=1}^{J} d(\bar{x}_j - \bar{x}_1) - \frac{1}{J + 1} \sum_{j=1}^{J+1} d(\bar{x}_j - \bar{x}_1) \right] - \gamma(1 - \theta) \frac{1}{J + 1} d(\bar{x}_{J+1} - \bar{x}_1)$$

or

$$\left( \pi(J + 1) - \pi(J) \right) + \gamma(1 - \theta) \frac{1}{J(J + 1)} \sum_{j=1}^{J} \left( d(\bar{x}_j - \bar{x}_1) - d(\bar{x}_{J+1} - \bar{x}_1) \right)$$

Now, since the agents are ranked in their distance to $\bar{x}_1$, the last term is (for each element of the sum) negative and gets more negative as $\gamma$ (or $(1 - \theta)$) increases. It thus follows that the gain from adding 1 or more agents decreases in $\gamma$ (or $(1 - \theta)$). This implies, by monotone comparative statics, that the optimal number of agents must decrease. Given that we always have an interval of agents by lemma 1, the radius $\tilde{d}$ must decrease. Notice (for later reference) that the agents who remain are all closer to $\bar{x}_1$ than the agent who drops out. ■
B Distance payoff functions

This appendix shows what kind of settings would generate a quadratic distance payoff function. Consider an organization that has a monopoly in $N$ different markets for a product with zero marginal cost. To sell in a market, the organization needs to send a sales agent who can independently set the price in his market. Price-setting is non-contractible. Let the demand in market $n$ be $Q_n = A - p_n$ with $p_n$ being the local price. The parameter $A$ is unknown, however, but each member of the organization has a subjective belief about its value. Let salesman $n$’s belief about $A$ be distributed according to $F_n$ with mean $\hat{A}_n$. He will then set $p_n = \frac{A_n}{2}$ so that the true profit for that market is

$$\frac{A\hat{A}_n}{2} - \frac{\hat{A}_n^2}{4} = \frac{A^2}{4} - \frac{1}{4}(A - \hat{A}_n)^2$$

If the organization ‘happens’ to have a fixed cost of $N(A^2 - 1) \frac{1}{4}$ then this gives a payoff function $\Pi = \frac{1}{4} - \frac{1}{4N} \sum_{n=1}^{N}(A - \hat{A}_n)^2$. In fact, any set of independent decisions that have the structure of an investment problem with return $\alpha I^2$ and cost $\beta I$ with uncertainty on $\beta$ will have a similar payoff structure.

For a distance payoff structure with the absolute value as distance measure, consider again an organization with a monopoly in $N$ different markets. In each market, demand is a unit mass at location $x \in \mathbb{R}$ with willingness to pay $V$. To sell in market $n$ the organization needs a local salesman who independently chooses the location of the local product $x_n \in \mathbb{R}$. Transportation costs for consumers are linear, so that the price they are willing to pay is $V - |x_n - x|$ as long as this is non-negative. Assume again that $x$ is unknown but that the agents have subjective beliefs (who are never more than $V$ wrong), then we get again a distance payoff function but now with $d(x, x_j) = |x - x_j|$. Again, any location problem of this type tends to give a similar payoff structure.
References


