COMPLEXITY AND ORGANIZATIONAL ARCHITECTURE

by

Yannis M. Ioannides[1]

March 20, 2011

Abstract

This paper examines how architectural features of organizations, that are made up of individuals screening projects, affect organizational screening performance. The paper explores consequences for organizational design of a theorem (attributed to Von Neumann by Moore and Shannon) on how to build reliable networks using unreliable components. The paper examines general properties of committee decision making and shows the superiority of committees with respect to composition, that is when each member of organization is replaced by a replica of the entire organization. The paper links with the modern Condorcet Jury Theorem literature. It also shows that screening performance is sigmoid in individual screening performance for hierarchies of polyarchies and for polyarchies of hierarchies. The supermodularity and sub-modularity properties of those structures allows us to link with results from the theory of teams. The screening performance is also sigmoid for a cognitive model that allows for individuals' own screening to be influenced by screening decisions of superiors and of subordinates. The paper examines the implications of such features for the limits to organizational performance.

JEL classification codes: D200, D230. Keywords: organizations, project evaluation, screening, organizational architecture, complexity, composition, social networks in firms, committees, reliability.

[1]Department of Economics, Tufts University, Medford, MA 02155, USA. `Yannis.Ioannides@Tufts.edu`
URL `http://www.tufts.edu/~yioannid/`.

`Ioannides Org Archi-Haller Conf Rev 03 20 11.tex`

COMPLEXITY AND ORGANIZATIONAL ARCHITECTURE[2]

1 Introduction

This paper revisits the literature on modelling organizations as networks of agents. Individual agents screen projects. Architectural features of organizations, that is how each agent's decision combines with those of others, affect the organization's screening performance. The paper emphasizes how multi-agent organizations may improve upon individual screening performance by suitable arrangement of the flow of decisions. The key result of the paper is grounded, in part, on a theorem of Von Neumann (1956) (see also Moore and Shannon (1956)) on how to build reliable networks using unreliable components. It is also conceptually related to the Condorcet Jury Theorem [Le Marquis de Condorcet (1785; 1994)]. The paper extends previous contributions by Sah and Stiglitz (1985; 1986; 1988) and by Ioannides (1987) by endogenizing organizational features and by demonstrating the advantages of combining different types of organizational architectures. The paper allows for more general decision making by means of cognitive models of screening performance.

Individual agents are engaged in screening projects. Architectural features of organizations, that is the specifics of how each agent's decisions combine with those of others, affect an organization's screening performance. For an organization's screening performance to improve over that of a single agent's, its screening function must be *sigmoid*, that is being originally convex and then becoming concave.[3] in individual performance, as measured by

[2]Support by the John D. and Catherine T. MacArthur Foundation through the Research Network on Social Interactions and Economic Outcomes and by the National Science Foundation under grant ACI-9873339 is gratefully acknowledged. I thank Kenneth Arrow, David Austen-Smith, Larry Blume, Toni Calvó-Armengol, Jacomo Corbo, Xavier Gabaix, Hans Haller, Semih Koray, Linda D. Loury, Glenn Loury, Hervé Moulin, Emmanuel Petrakis, Sudipta Sarangi, Dimitris Vayanos, Bauke Visser and other participants at presentations at various venues and at the Conference in the Honor of Hans Haller, Louisiana State University, February 25–26, 2011. I am solely responsible for any errors.

[3]A sigmoid function, as in an $\mathcal{S}$ shape, is well represented by the logistic function.

the probability that the organization accepts (rejects) a good (bad) project. This is, indeed, the case for organizations with *mixed* Sah–Stiglitz architectures, such as hierarchies made up of components that are themselves polyarchies and polyarchies made up of components that are themselves hierarchies, but not for pure Sah–Stiglitz architectures. This property is in turn critical for endogenizing individual screening performance. The model is extended by means of cognitive features that also allow for individuals' own screening outcomes to combine with influence from subordinates and superiors. The paper examines the implications of cognitive features for the limits to organizational performance of complex organizational architectures with top to bottom feedback.

A strand of the current research on organizations invokes the notion of organizational *complexity*. In computer science, *computational* complexity pertains to how hard it is to perform computations needed to solve computationally well-defined problems [Papadimitriou (1994)]. Recent uses of the term organizational *complexity* are not as well-defined. Some authors appeal to Simon (1962), who emphasizes that the hierarchical nature of complex systems in nature serves an important purpose. That is, hierarchical structure isolates the impact of disturbances and lessens the effects of shocks and mistakes. The notion of complexity in this paper is similar to that articulated by Sah and Stiglitz (1988), p. 467–468, and to that developed by Mount and Reiter (1998; 2002), Rivkin (2001), and especially Visser (2001a). The latter defines as *complex* organizational structures that are made up of a "large number of divisions or hierarchical layers or if they contain many interdependent parts the individual functioning of which is of importance to the overall performance of the organization" [Visser (2001a), p. 1].[4]

[4] Visser (2001a) defines *complexity* as the level of detail that is necessary to correctly assign agents to positions in the organizational structure; *performance* in terms of the maximum expected profit associated with the structure; and, *robustness* in terms of the maximal extent in which expected profits on implemented projects fall short of the optimal case which would obtain if agents had been correctly assigned. Visser (2001a) shows that increasing organizational complexity is associated with reducing robustness, at least for the specific organizational forms which he examines in the paper. Visser (2001b) argues in favor of cognition aspects of complexity. He proposes to study organizations in terms of three possible types of information about agents' characteristics that may be used in designing organizations, that is, no information, ordinal information, and cardinal information. Fioretti and Visser (2002) focus away from defining organizational

The present paper associates complexity with organizational architectures that are made up of divisions that are organized in a particular way, while each of them is made up of agents organized in a different way. Such organizations are of course quite common, and yet have not been studied extensively. As we see in further detail below, it is an important result of the paper that it is such "mixing" of organizational architectures that makes it possible for an organization of many agents to perform screening that is superior to that by an individual.

The remainder of this paper is organized as follows. Section 2 defines firms as networks of interdependent decision makers, elaborates on architectural features of organizations, and motivates the principle of composition by recalling the von Neumann–Moore–Shannon theorem of optimal network design. Section 3 endogenizes individual screening performance via setting the wage rate so as to maximize profit and studies specific instances of complex organizations that involve both hierarchical and polyarchical features. Section 4 turns to a model of organizations, whose agents' screening performance is described in terms of a cognitive model. This allows us to account for uncertainties affecting an agent's evaluation, that are inherent in decision making, and for the influence from co-workers, and poses the problem of optimizing organizational architecture.

2 The Firm as a Network of Interdependent Decision Makers

This paper models firms as networks of interdependent decisions, where each project receives a "thumbs up" or "thumbs down" from each individual screening agent. The problem of organizational architecture is to determine how these individual decisions are aggregated into a final organization-wide choice, that is whether a particular project should be adopted or rejected.

---

complexity directly in terms of organizational architecture and emphasize instead the importance of linking with the human cognition of "a structure or behavior."

2.1 Individual Screening Performance and the Objective of the Firm

A project's quality is unobservable and may be either good or bad, with probability α or $1-\alpha$, respectively, where $0 < \alpha < 1$, is a parameter for a population of projects, from which workers draw randomly. Screening is imperfect but informative. That is, a good project is more likely to be accepted than a bad project. A firm may influence its workers' screening effectiveness via the wage rate. We assume that wages are paid before agents make screening decisions and before project quality is revealed.

Let the set $\mathcal{I} = \{i : i = 1, \ldots, I\}$ represent the members of an organization. The organizations, that is the firm's costs are only labor costs, $\sum_{i \in \mathcal{I}} w_i$, with agent i receiving a wage rate w_i. Agents are paid before their screening is completed. Greater remuneration increases the effectiveness of screening. As a function of the wage rate, the probability of an agent's accepting a good project, $p_g = P_g(w)$, $P_g : R_+ \to [0, 1]$, is increasing and concave in w, the wage rate. The probability of an agent's rejecting a bad project, $1 - p_b = 1 - P_b(w)$, $P_b : R_+ \to [0, 1]$, is increasing and concave in the wage rate. Functions $P_g(w), P_b(w)$ are assumed to be twice differentiable. All these assumptions make the probability of accepting a good project and of rejecting a bad one increasing and concave in the wage rate. Assumptions

$$0 < P_b(w) < P_g(w) < 1 \tag{1}$$

ensure that screening is informative in the sense that a good project has a higher probability of being accepted than a bad one. One may interpret $1 - p_g$, the probability that a good project be rejected, as a Type-I error, and p_b, the probability that a bad project be accepted, as the Type-II error, for an individual agent.

Let $R_g > 0$ and $R_b < 0$ denote the revenue to an organization from accepting a good and bad project, respectively. The probability that a firm accept a project, conditional on its quality $j = g, b$, is represented by the organization's *screening function* $\mathcal{P}(p_j)$. An organization's screening function $\mathcal{P}(p)$, $\mathcal{P} : [0, 1] \to [0, 1]$, is the probability that the organization accept a project, as a function of p, the probability that each member accept a project. Function $\mathcal{P}(p)$ summarizes the impact of organizational architecture on screening performance.

The expected profit from considering a randomly selected project is given by

$$R_g \cdot \alpha \cdot \mathcal{P}(p_g) + R_b \cdot (1-\alpha) \cdot \mathcal{P}(p_b) - \sum_{i \in \mathcal{I}} w_i. \tag{2}$$

We return further below in section 3.5 to the problem of endogenizing individual screening performance. Next we take up the determination of an organizations's screening function under different organizational architectures.

2.2 Organizational Architecture and the Principle of Composition

Sah and Stiglitz [Sah and Stiglitz (1985; 1986; 1988), and Sah (1991)] define *polyarchies* as organizational structures, *architectures*, in their terminology, with agents operating "in parallel:" a polyarchy approves of a project if at *least one* of the agents in the organization approves it. For example, the screening function for a *polyarchy* of two agents is given by $\mathcal{P}_P(p) = 1-(1-p)^2 = 2p - p^2$ and is increasing and concave in p. Similarly, they define *hierarchies* as organizations with agents operating "in series": a hierarchy approves of a project only if *all* agents approve it. For example, the screening function for a *hierarchy* with two agents is given by: $\mathcal{P}_H(p) = p^2$, and increasing and convex in p. Clearly, both polyarchies and hierarchies are informative in that a good project is more likely to be accepted than a bad project.

Some properties of these stylized architectures readily follow from their definitions. The fact that $\mathcal{P}_P(p) > \mathcal{P}_H(p)$, $\forall p \in [0,1]$ implies that polyarchies accept more good and more bad projects than hierarchies. Equivalently, since $1-\mathcal{P}_P(p) < 1-\mathcal{P}_H(p)$, hierarchies reject more good and bad projects than polyarchies.

Sah and Stiglitz (1988) apply this framework to study of committee decision making. E. g., a polyarchy is a committee that operates with a plurality of 1, and a hierarchy with unanimity. Sah and Stiglitz do discuss briefly the possibility of more complex organizations, such as committees of committees and polyarchies including hierarchies as their elements and vice versa [*ibid.,* 467–468] and the possibility of arbitrary improvement, as the number of agents become infinitely large in organizations made up of committees which are themselves constituent elements of other committees. However, they do not recognize the particularly attractive properties of cross composition, which are explored in the present paper.

Ioannides (1987) relies on von Neumann (1956) and Moore and Shannon (1956) to show how it is possible, in general, to improve the screening performance of organizations by means of *complicating* their architecture in the very precise sense of *composition*. Composition, a term originated by Moore and Shannon (1956), p. 197, involves replacing each agent with a replica of the entire organization, and provided that the architecture satisfies certain key properties, it is possible to design an organization that satisfies given performance criteria. That is, the organization must reject projects, whose probability of being accepted by an agent is below a specified threshold, and must accept projects whose probability of being accepted by an agent is above a specified threshold.[5] Ioannides (1987) shows that pure Sah–Stiglitz architectures are *not* improvable by means of composition.

Although *composition* may make organizations more reliable, it does so at great increase[6] in organization size. That is, replacing each member of an organization of size n with n members causes the size to increase from n to n^2, and the next composition increases it to n^4, and to n^{16}, and so on.[7] This renders composition of theoretical rather than practical

[5]Moore and Shannon (1956) in turn credit Von Neumann (1956) for the original idea of using (unreliable) *Sheffer stroke organs* as components of reliable organisms. A Sheffer stroke is the Boolean operation "(`.not.` A) `.and.` (`.not.` B) on Boolean variables A and B. It has the property, which Von Neumann exploits, that all logical operations can be generated from it. A Sheffer stroke organ is the logical operation with two binary inputs A and B which performs this logical operation, possibly unreliably. The von Neumann approach is noteworthy because it deals explicitly with the stochastics of error-prone components.

[6]Moore and Shannon are interested in modelling engineering systems, which are not "intelligent" and have to satisfy strict performance standards, that is θ_1 and θ_2 are very small numbers. This causes the number of agents to increase sharply with the pre-specified performance standards. The sensitivity of organizational size to the stated performance requirements for an organization is characterized by bounds obtained by Moore and Shannon (1956).

[7]Note that this is much faster than the increase in the number of agents when the number of levels of a pyramidal hierarchy, K, while the span of control s remains fixed. In the latter case, the number of agents in a hierarchy of K levels with a span of control of s, $1 + s + \ldots + s^K$ increases by s^{K+1}, when the number of levels increases from K to $K + 1$.

significance, as a way of improving an organization's screening performance.[8] It is also the reason why Sah and Stiglitz do not consider it in depth.

For an organizational architecture to lend itself to improvement by composition its screening function must have a single fixed point in $(0, 1)$ and must cut the 45^o line from below at that point. In other words, since the 45^o line represents the response of an organization with a single agent, there must exist a value of the screening probability between 0 and 1 that makes an individual agent be equivalent to an entire organization: $\mathcal{P}(p) = p$. A round of composition reduces the probability of a project's being accepted by the organization, if all agents' acceptance probability is below this characteristic value, and increases it, if it is above. See Figure 1. In Figure 1, starting with an organization with screening function $\mathcal{P}(p)$, composition associates with individual screening performance p organizational performance $\mathcal{P}(\mathcal{P}(p))$. Stated in the form of a proposition, we have the following. All proofs are in the Appendix.

<u>Proposition 1.</u> *An organization's screening function $\mathcal{P}(p)$ that possesses a single fixed point p^*, $p^* \in (0, 1)$, satisfies $\frac{d}{dp}\mathcal{P}(p^*) > 1$, (the screening function cuts the 45^o line from below at the fixed point) and is therefore more likely to accept good projects and to reject bad projects than any individual agent, provided that $p_b < p^* < p_g$.*

The mathematical intuition of this result should be straightforward to economists familiar with dynamical systems. The higher-order iterates of a time map share with the original map the same fixed points. Similarly, composition is the "spatial" counterpart for the screening function of iterating the time map of a dynamical system. The fixed point of the screening function reflects fundamental aspects of organizational architecture and is invariant to "spatially" iterating the screening function.[9] Composition would not improve an organization's

[8] Referring to this process as increasing organizational *complexity*, would find Mount and Reiter (1998) in agreement. We return tho this further below.

[9] We note that, however tempting, a direct comparison of the von Neumann–Moore–Shannon theorem to results on neural networks as universal approximators would be simplistic. The von Neumann–Moore–Shannon theorem exploits the particular nonlinearity of the probability that an organization will approve of a project as a function of the individual performance so as to satisfy desired performance standards by means of spatially iterating that function. The universal approximator theorem for neural networks [Haykin

screening performance over that of an individual unless $p_b < p^* < p_g$, a condition that links screening characteristics of individuals to organizational architecture. Finally, by Theorem 3, Harrison (1965), p. 243, by iterating composition, screening is arbitrarily improved:

$$\lim_{k\to\infty} \mathcal{P}(\mathcal{P}(\mathcal{P})\ldots) = \begin{cases} 1 & p > p^*; \\ p^* & p = p^*; \\ 0 & p < p^*. \end{cases}$$

The von Neumann–Moore–Shannon theorem prompts us to explore how complicating organizational architecture by means of what we refer to as *cross*-composition may improve organizational screening performance with a slower increase in the number of agents than what composition requires. We see shortly that cross-composing a hierarchy with a polyarchy (and vice versa) just *once,* that is, by replacing each agent in a pure hierarchy by a component which itself is a polyarchy, makes the organization's screening function sigmoid. Therefore, it is, in principle, possible to make an organization more effective both in screening good and bad projects. For example, cross-composition of a hierarchy with a polyarchy by having polyarchies as constituent elements of a hierarchy combines the desirable features of both types of pure Sah-Stiglitz architectures. That is, the resulting organization will likely reject more bad projects and approve more good projects. We turn next to exploring further the concept of cross-composition. It is interesting to do so because, in spite of the arguments of Sah and Stiglitz that most real-life organizations are not complex, organizations of minimal complexity do exist. For example, U.S. academic institutions are often polyarchical at lower levels, as where different disciplines make recommendations regarding recruitment in terms of their own criteria, but hierarchical at higher levels, where deans and the provost are ultimate arbiters of academic recruitment decisions. Also, some scientific journals use referees as hierarchies while editorial boards function as polyarchies.

A brief excursion into related literature is in order. Mount and Reiter (1998; 2002) and Reiter (1996) develop a model of the firm, in which organizational structure emerges as a solution of an optimization problem. The problem involves how to carry out computations and

(1999), 208–209, 229] exploits the particular nonlinearity of the sigmoid function that makes it perform well as a functional form for the components of a power series.

trades off the "complexity" of economic computations in a given class of economic environments against constraints, which trade off limitations on the abilities of agents to compute and communicate. Computations are decomposed in terms of "auxiliary" computations that can be expressed as a *superposition* (composition, in von Neumann's terminology, which we adopt) of the primitive functions performed in terms of "lower level" computations (or agents). Their measure of complexity is then simply the depth of superposition, which in our case corresponds to the number of iterations, which correspond to the number of hierarchies. The model of the present paper is conceptually similar to the Mount–Reiter model of the firm. Our agents' evaluations are akin to computations, and organizational architecture is akin to the execution of computations in the Mount and Reiter model.

2.3 How about Committees?

One may wonder about the properties of decision-making by committees. Intuitively, since hierarchies and polyarchies are extreme special cases of committees, the general case of committees ought to combine the properties of the two extremes. [10] Consider a committee of K members, each evaluating a project independently of one another and approving it with probability $p = p_g$, if it is good, and $p = p_g$, if it is bad; the probability that it is approved by $k \leq K$ members is given by the probability of k successes in K independent Bernoulli trials: $\frac{K!}{k!(K-k)!}p^k(1-p)^{K-k}$. The probability that is approved by more than k members, $\mathcal{P}_B(p; k+1, K)$, is given by:

$$\mathcal{P}_B(p; k+1, K) = \sum_{n=k+1}^{K} \frac{K!}{n!(K-n)!} p^n (1-p)^{K-n} = \frac{\mathcal{B}(p; k+1, K-k)}{\mathcal{B}(k+1, K-k)} = \frac{\int_0^p t^k (1-t)^{K-k-1} dt}{\int_0^1 t^k (1-t)^{K-k-1} dt}, \tag{3}$$

where $\mathcal{B}(p; k+1, K-k), \mathcal{B}(k+1, K-k)$ denote the incomplete beta and complete beta functions, respectively, and the last equality rests on the integral representation of the beta function. The properties of $\mathcal{P}_B(p; k+1, K)$ are summarized in the following proposition.

Proposition 2.

[10] Toni Calvó-Armengol made this point during a conversation in Cambridge, Massachusetts, November 2006.

For a committee of K members, each deciding independently with probability p, a project is approved by more than k members with probability $\mathcal{P}_B(p; k+1, K)$, given by (3).

Part A. *Probability $\mathcal{P}_B(p; k+1, K)$ is increasing in p and concave (convex), if $\frac{k}{K+1} < (>)\ p$, and satisfies the conditions of Proposition 1. It follows that a committee's approval of a project by majority voting is in general sigmoid in p.*

Part B. *Probability $\mathcal{P}_B(p; k+1, K)$ is increasing with decreasing increments in K, for large K, if $p > \frac{1}{2}$; it is increasing with increasing increments in K, for small K, if $p > \frac{1}{2}$; it is decreasing with increasing increments in K, if $p < \frac{1}{2}$.*

The fact that the probability that a committee approve of a project by majority voting is *sigmoid* in p is the single most important feature of committee decision-making for our purposes. It follows from Proposition 1 that organizations made up of committees may be improved upon by composition. This supports the wisdom of committees' having subcommittees of their members, an instance of composition. A committee is more likely to approve of a project than a single member, for good projects, and less likely, for bad projects, provided that the individual probability of approval is large enough as compared to a (slightly weakened) plurality requirement in percentage terms, $\frac{k}{K+1} < p$. Correspondingly, the probability of a wrong decision by the committee is decreased by composition if the individual probability of disapproval is low enough. The present paper provides a contribution over and above Sah and Stiglitz (1988), p. 468 who do acknowledge the power of composition, by highlighting the mathematical underpinnings of how composition improves committee decision making. The results of Part A are a bit surprising because as the number of Bernoulli trials increases, each of which occurs independently and with probability p so do both the mean *and* variance, $(Kp, Kp(1-p))$, and therefore, the probability distribution becomes more spread out.

Proposition 2 also helps link our approach with the Condorcet Jury Theorem and it does generalize it in the sense that elaborates on the outcome under different plurality rules. According to the classical statement of the theorem, if $2N+1$ jurors act independently, each with probability p exceeding one-half of making the correct decision, then the probability that the jury, deciding by majority voting, makes the correct decision increases and tends to

1, as N tends to infinity.[11] Therefore, just as with organizations it would be better to entrust a decision to a group of individuals of lesser competence than to a single individual of greater competence [Boland (1989)]. There are aspects of the Condorcet Jury Theorem that attract current interest. Most notably, Austen-Smith and Banks (1996) show that the theorem is not immune to insincere behavior, even if individuals have a common preference for selecting the better alternative. Koriyama and Szentes (2007) return to the basics of group decision making in the light of information acquisition and find that the optimal committee size is finite.[12]

2.3.1 The Superiority of Committees

What can one say about decision-making mechanisms where the collective decision rests on a plurality rule, like in committee voting? The little known theory of quorum functions allows us to state that committees are the "best" possible mechanisms for aggregating group preferences. Relying on Moore and Shannon (1956) and Harrison (1965), I have the following. A Boolean function of n variables $(x_1, \dots, x_n)$ is said to be a *quorum function* if there exists q, such that $0 \leq q \leq n$, and:

$$f_q(x_1, \dots, x_n) = \begin{cases} 0, & \text{if less than } q \text{ variables are equal to 1;} \\ 1, & \text{if more than } q \text{ variables are equal to 1.} \end{cases}$$

Proposition 3 merely restates in my setting Theorem 2, Moore and Shannon (1956), p. 202–2003 [see also Harrison (1965), p. 246, and is provided without proof.

Proposition 3. *If $f_q(p)$ denotes a quorum function for $x_1 = x_2 = \ldots = x_n = p$, that intersects with the 45^o line at the same point as the screening function of an organization with n agents, $\mathcal{P}$, p^*, then:*

[11] This readily follows from Proposition 2, especially Equ. (22) – see Appendix. Committee approval tends to 1, when K tends to infinity, if $p > \frac{1}{2}$, and to zero, otherwise.

[12] We note that although the Condorcet Jury Theorem is conceptually related to the von Neumann and Moore–Shannon results, no such link has been made (to the best of my knowledge) by either the von Neumann and Moore–Shannon literatures, or the modern Condorcet Jury Theorem literature.

Part A.

$$\frac{d\mathcal{P}(p)}{dp}|_{p=p^*} < \frac{df_q(p)}{dp}|_{p=p^*}.$$

Part B.

$$\begin{array}{ll} \mathcal{P}(p) > f_q(p) & 0 < p < p^*; \\ \mathcal{P}(p) < f_q(p) & p^* < p < 1. \end{array}.$$

Proposition 3 implies that any arbitrary organizational architecture is dominated, from the specific viewpoint of its screening performance and how it may be improved by composition, by committee-type architectures, where the decision is made with a plurality rule.

3 Sah–Stiglitz Architectures as Building Blocks of Complex Organizations

Proposition 4 below summarizes the screening properties, for both for good and bad projects, of pure Sah–Stiglitz organizations made up of many agents with identical screening performance. Its proof is elementary and therefore deleted.

Proposition 4.

Part A. *The screening function of a hierarchy of* I_H *members* $\mathcal{P}_H(p) = p^{I_H}$, *is convex increasing in the individual screening probability* p, *convex decreasing in the number of agents* I_H, *and satisfies* $\mathcal{P}_H(0) = 0$, $\mathcal{P}_H(1) = 1$. *For good projects,* $\mathcal{P}_H(p)$ *is a concave (convex) increasing function of the wage rate, provided that* $I_H - 1 < -\frac{P_g''/P_g'}{P_g'/P_g}$ $(>)$ 0; *for bad projects, it is decreasing convex function of the wage rate.*

Part B. *The screening function of a polyarchy of* I_P *members,* $\mathcal{P}_P(p) = 1 - (1-p)^{I_P}$, *is concave increasing in* p, *concave increasing in* I_P, *and satisfies* $\mathcal{P}_P(0) =$, $\mathcal{P}_P(1) = 1$. *For good projects,* $\mathcal{P}_P(p)$ *is a concave increasing function of the wage rate; for bad projects, it is convex (concave) decreasing, provided that* $\frac{P_b''/P_b'}{P_b'/1-P_b} > I_P - 1$ $(<)$ 0.

We discuss further below the case of endogenizing screening performance by means of the wage rate. It would thus be helpful to note that if the probability of accepting good projects and of rejecting bad projects are, roughly speaking, sufficiently insensitive to the wage rate, maximization of expected profit defines an optimal wage. An optimal size exists for the polyarchy, but for the hierarchy, expected revenue is decreasing convex in size and so is expected profit.

Complicating simple Sah–Stiglitz architectures by means of composition preserves the mixing and sorting properties, because a concave (convex) function of concave (convex) functions is concave (convex). However, it is not particularly interesting to examine such compositions because the simple Sah–Stiglitz architectures do not satisfy the Moore-Shannon criterion, and therefore, such compositions do not improve an organization's screening performance [Ioannides, *op. cit.*]. It is for this reason that we turn next to *cross*-composition with pure Sah-Stiglitz architectures.

3.1 Cross-Composition: Complex Architectures with Sah–Stiglitz Components

As Sah and Stiglitz emphasized in their original contribution, pure Sah–Stiglitz architectures make different kinds of errors, with polyarchies accepting more good *and* bad projects and hierarchies rejecting more good *and* bad projects. It thus appears natural to explore complex organizational architectures that by combining pure architectures may enhance their desirable properties and weaken their undesirable ones, that is make it more likely for the organization to accept good projects and to reject bad ones. This is exactly what *cross-composition* of pure architectures makes possible. In a nutshell, such minimally complex organizations exhibit screening performance that is a sigmoid function of individual screening performance and under certain conditions satisfies one of the conditions of Proposition 1, namely that it possesses a unique fixed point in $(0,1)$ where it intersects the 45^o line from below. We see below that for all organizations that are formed by cross-composition of pure hierarchies and polyarchies, there exists a specific value of individual screening probability, below (above) which, the probability of accepting a project by the organization is less (greater) than the probability that an individual agent approves a project. Therefore, as long as screening is informative, organizational design can improve, in principle, on individual screening effectiveness.

For organizations with a total number of $I = |\mathcal{I}|$ agents, we let $I_1, I_2, \ldots, I_K$, agents be placed at hierarchy levels $1, 2, \ldots, K$, respectively, where all agents at each hierarchy level k have identical screening performance p_k, but agents across different hierarchy levels may differ, $p_k \neq p_j$. Let the set of members $\mathcal{I}_k$ at level k be organized as a polyarchy. Then the screening function of a hierarchy of K polyarchies, that is, the probability of accepting a project by such an organization, is given by

$$\mathcal{P}_{HP}(p_1, I_1; \ldots, p_K, I_K) = \prod_{k=1}^{K} [1 - (1 - p_k)^{I_k}], \tag{4}$$

where we adopt a mnemonic rule that HP stands for *h*ierarchies of *p*olyarchies.

Working in like manner, consider a polyarchy consisting of $\frac{I}{K}$ hierarchies, each made up of K members. Across all hierarchies, agents at the same level k have equal screening performance p_k. Then the probability of accepting a project by such a polyarchy of hierarchies, its screening function, is given by

$$\mathcal{P}_{PH}(p_1, \ldots, p_K; K, I) = 1 - \left(1 - \prod_{k=1}^{K} p_k\right)^{\frac{I}{K}}. \tag{5}$$

Proposition 5 and 6 below state some general properties of these minimally complex alternative organizational architectures, polyarchies of hierarchies and hierarchies of polyarchies, whose agents have identical screening performance: $p_k = p, \forall k$. Below in Section 3.5 we examine conditions under which such a feature is indeed optimal. Throughout, we let the total number of agents I be sufficiently larger than the number of hierarchies K, so that $\frac{I}{K}$ may be treated as an integer.

3.2 Hierarchies of Polyarchies

Proposition 5.

The screening function of a hierarchy of polyarchies, with a total number of I members who are arranged in a hierarchy of K polyarchies, $I > K \geq 1$, each of size $\frac{I}{K}$, given by (4), *with* $p_k = p, \forall k$,

$$\mathcal{P}_{HP} \equiv \mathcal{P}_H \left[\mathcal{P}_P(p)\right] = \left(1 - (1-p)^{\frac{I}{K}}\right)^K , \tag{6}$$

is monotone increasing in $p \in (0,1)$*, satisfies* $\mathcal{P}_{HP}(0) = 0,\ \mathcal{P}_{HP}(1) = 1$*, has a unique fixed point* p^*_{HP}*, and a unique inflection point* $\tilde{p}_{HP} = 1 - \left(\frac{I-K}{K(I-1)}\right)^{\frac{K}{I}}$ *, both in* $(0,1)$*. It is increasing and concave in* I*, decreasing and convex in* K*, is sigmoid with respect to* $p \in (0,1)$ *and intersects the* 45^o *line from below at its unique fixed point* p^*_{HP}.

As more hierarchy levels are added, the inflection point of the screening function increases making screening more demanding; as more agents are added while holding the number of hierarchy levels constant, the fixed point of the screening function decreases. It is therefore not possible both to improve the organization's performance in rejecting poor projects and in approving good projects in this fashion. A tradeoff emerges between overall probability of acceptance and effectiveness of screening performance. Increasing the number of hierarchy levels reduces the probability that an organization accept any given project. But if the probability of acceptance is high, such an increase makes it more likely that better projects would be accepted.

3.3 Polyarchies of Hierarchies

The properties of a polyarchy of I agents that is made up of $\frac{I}{K}$ hierarchies, each of which has K agents, is summarized by the following proposition.

Proposition 6.

The screening function of a polyarchy with a total number of I *members, which is made up of* $\frac{I}{K}$ *hierarchies each with* K *levels, given by (5), with* $p_k = p, \forall k$,

$$\mathcal{P}_{PH}(p) \equiv \mathcal{P}_P \left[\mathcal{P}_H(p)\right] = 1 - \left(1 - p^K\right)^{\frac{I}{K}} , \tag{7}$$

is monotone increasing in $p \in (0,1)$*, has an unique fixed point* p^*_{PH}*, and an inflection point* $\tilde{p}_{PH} = 1 - \left(\frac{K-1}{I-1}\right)^{\frac{1}{K}}$ *, both in* $(0,1)$*, provided that* $I > K > 1$*. It is increasing concave in the total number of agents* I *and is convex decreasing in the number of hierarchy levels* K*, is sigmoid with respect to* $p \in (0,1)$*, and intersects the* 45^o *line from below at its unique fixed* *point* p^*_{PH}.

We note that inflection point $\tilde{p}_{PH}$ is decreasing in the size of the organization but increasing in the number of hierarchies. A tradeoff emerges between overall probability of acceptance and effectiveness of screening performance, the counterpart to the one for a hierarchy of polyarchies: increasing the number of hierarchy levels reduces the probability that an organization accept any given project while favoring better projects.

3.4 Traditional Hierarchical Architecture with a Fixed Span of Control

We saw that both hierarchies of polyarchies and polyarchies of hierarchies with identical agents are associated with screening performance functions that are sigmoid in individual screening performance. Clearly, it is the mixing of architectures that is responsible for the sigmoidicity of the performance function and thus for the possibility of improvements similar to those enabled by composition. However, even cross-composing the pure architectures just once moves the organization toward improved screening performance.

This conclusion is bolstered by the following observation.[13] Consider what is arguably the standard architecture of hierarchical organizations with a hierarchy of K levels and fixed span of control, s. That is, there is a single agent at level $k = 1$, the top, to whom s agents at level $k = 2$ report. Reporting to each agent at level $k = 2$ are s agents, at level $k = 3$, and so on until level $k = K$, where there are s^{K-1} agents. The organization is thus made up of $1 + s + s^2 + \ldots + s^{K-1} = \frac{s^K - 1}{s-1}$ agents. We assume that all agents have identical screening performance p and consider first the screening function of a two-level organization. The probability that the agent at level $k = 1$ receive a positive recommendation by her subordinates is given $1 - (1-p)^s$ and the organization's screening function is $\mathcal{P}_2(p) = p\left(1 - (1-p)^s\right)$. This function has the standard properties, that is it is strictly increasing in $p, p \in (0,1)$, with $\mathcal{P}_2(0) = 0,\ \mathcal{P}_2(1) = 1$. Furthermore, $\mathcal{P}_2(p)$ has no fixed point in $(0,1)$, $\mathcal{P}_2'(0) = 0$, and $\mathcal{P}_2'(1) = 1$. Its inflection point is at $\tilde{p} = \frac{2}{1+s}$, at which $\mathcal{P}_2$ lies below the 45^o

[13] I thank Hervé Moulin for suggesting that i explore this point.

degree line and its slope is equal to $1+\left(\frac{s-1}{s+1}\right)^{s-1}$. It is an open question whether these basic properties are satisfied by the screening function for $K-$level hierarchies. Since it true that $\mathcal{P}_K(p) < p$, the screening performance does not lend itself to improvement by increasing the number of agents.[14]

In sum, traditional hierarchical organizations (whose architectures have the graph topology of trees) in an intuitive sense also combine the properties of pure hierarchies and polyarchies. That is, in a fashion resembling that of cross-composition, agents who report to the same superior act as in parallel (they act as polyarchies), whereas agents to whom other agents report and who belong to the same line of authority act as in series (they act as hierarchies). It thus interesting that traditional tree-like architectures fall short of enhancing their desirable properties of one type or architecture and weakening their undesirable ones, when the number of agents increases, in the way that cross-composition does.

3.5 Endogenous Individual Screening Effectiveness

Next we endogenize individual screening performance by optimally setting the wage rate for an organization with a given screening function $\mathcal{P}[\cdot]$. A higher wage sharpens the screening effectiveness of the typical agent in screening both good and bad projects. Recall our assumption that as a function of the wage rate the probability of an individual's accepting a good project, $P_g(w)$, is increasing concave, and the probability of accepting a bad project, $P_b(w)$, is decreasing convex. Working from the definition of expected profit (2), after we allow for the fact that in general, the members of the same group $\mathcal{I}_k$, of size I_k, earn the same wage, we have:

$$\sum_{k=1}^{K} I_k w_k. \tag{8}$$

[14]This question may be pursued further by induction because the properties of $\mathcal{P}_K(p)$ may be investigated iteratively in terms of those of $\mathcal{P}_{K-1}(p)$, since:

$$\mathcal{P}_K(p) = p\left[1 - \left[1 - \mathcal{P}_{K-1}(p)\right]^s\right].$$

The wage cost is simply Iw, if all agents are identical. The proposition that follows reveals a critical role of the fact that the screening function is sigmoid when screening performance is endogenized via the wage rate.

Proposition 7.

A necessary condition for the existence of an optimal screening effectiveness in an organization composed of identical agents is that the organization's screening function be sigmoid in its argument.

Two remarks are in order. First, for the optimal wage w_{opt} to also cause an organization consisting of identical agents to be more effective in screening projects than an individual agent, it must be the case that the probability of accepting a bad project should be below p^*, the fixed point of $\mathcal{P}[\cdot]$, and the probability of accepting a good project should be above it, or

$$P_b(w_{opt}) < p^* < P_g(w_{opt}). \tag{9}$$

Practically speaking, the nearer that the fixed and the inflection points, $\hat{p}$ and $\tilde{p}$, of the organization's screening function are to one another, the more likely it is that condition (9) be satisfied.

Second, the larger is the number of agents, the smaller the optimal wage rate is likely to be. This follows from (23) using a qualitative argument. The second derivative of the expected profit will be decreasing in w for high values, in the canonical case of a sigmoid organizational screening performance function. Then the larger is I, the smaller w_{opt} would be.

The potential complexity of an organization's chart invites the obvious question: would it pay for a firm to differentiate agents in terms of their screening performance? If agents' screening responds to incentives, then it may pay the firm to differentiate workers depending upon their specific location within the organization. Our results are summarized in the proposition that follows.

Proposition 8.

For a hierarchy of K levels, each of which is a polyarchy with $I_k > 1, k = 1, \ldots, K$

members, whose screening function is given by (4), we have that for any two polyarchies at levels k_1, k_2 :

Part A. $w_{opt,1} \leq w_{opt,2}$, *provided that* $I_{k_1} \geq I_{k_2}$.

Part B. *Optimal wages are equal for all agents and numbers of agents at each hierarchy level are all equal to* $\frac{I}{K}$.

In other words, our result is that given the number of agents at different hierarchical levels, agents at different levels are paid more when they are fewer in number. It is thus natural to think of agents who are paid more as occupying higher levels of a "pyramidal" hierarchy.

3.6 Sorting versus Mixing with Heterogeneous Agents and Complex Architectures

Agents may differ in terms of screening effectiveness, how should agents be grouped across firms in terms of their screening effectiveness?[15] Let us consider two screening organizations, each of which employs two agents, for simplicity, and chooses them out of a total of four agents. There are two of each of two different types of agents, with screening effectiveness p_1 and p_2, respectively. Each firm's output is proportional to a firm's probability of approving a project.

3.6.1 Hierarchies and Sorting

If both organizations are organized as hierarchies, then each organization's expected revenue is proportional to the product of the screening probabilities of its workers, and thus a convex function of their p's, provided that such screening by each of the two workers is independent. Efficiency then dictates that agents working for hierarchical organizations be *sorted.* That is, total expected revenue by two organizations employing a total of 4 agents that are organized as hierarchies is greater if both agents of the same type work together, $p_1^2 + p_2^2$, than if they

[15] This problem is reminiscent of segregation by skill, discussed by Kremer and Maskin (1996).

work separately, $2p_1p_2$:

$$p_1^2 + p_2^2 \geq 2p_1p_2.$$

This result generalizes readily to the case of I^2 agents. That is, for I identical hierarchies, each employing one of each type of agent, $p_i, i = 1, \ldots, I$,

$$p_1^I + p_2^I + \ldots + p_I^I \geq I p_1 \cdot p_2 \cdots p_I.$$

These results carry over to the comparison of expected profits, provided that $R_g > -R_b$ and that agents are paid the same. Therefore, depending upon the sensitivity of individual screening effectiveness to the wage rate, the possibility exists for this sorting result to lead to unequal wages. It is possible to show that more effective workers would be paid more, independently of the wage rate for less effective workers.

3.6.2 Polyarchies and Mixing

If, on the other hand, both organizations are organized as polyarchies, then expected revenue is proportional to a concave function of screening probabilities. Efficiency dictates that agents be *mixed.* That is, the concavity properties of the polyarchy discussed above imply that total expected revenue by two organizations that are organized as polyarchies is greater if each organization employs agents of different type than if each organization employs agents of the same type,

$$2(1 - (1 - p_1)(1 - p_2)) \geq 1 - (1 - p_1)^2 + 1 - (1 - p_2)^2.$$

This result that hierarchies favor *mixing* of agents generalizes readily to the case of many agents. It also carries over to a comparison of expected profits, provided $R_g > -R_b$ and that agents are paid the same. It is possible to show that more effective agents would be paid more than less effective agents but the determination of the wage structure reflects, not surprisingly, both screening functions.

3.6.3 Sorting versus Mixing with Heterogeneous Agents and Complex Architectures

What can we say about the properties of organizations that are hierarchies of polyarchies or polyarchies of hierarchies that are made up of heterogeneous agents? When does it pay to mix different types of agents in the same organizations and when does it pay to sort them? We consider first the case where I agents of type 1 and I agents of type 2 are available 2 firms, and assume the number of agents of each type to be even. If both firms are hierarchies of polyarchies, then expected output from two identical firms, both of which consists of $\frac{I}{2}$ agents of each type, is less than the sum of expected outputs of two firms, one of which of which is made up of $\frac{I}{2}$ identical agents of one type and the other $\frac{I}{2}$ of the other type:

$$2\left(1-(1-p_1)^{\frac{I}{2}}\right)\left(1-(1-p_2)^{\frac{I}{2}}\right) \leq \left(1-(1-p_1)^{\frac{I}{2}}\right)^2 + \left(1-(1-p_2)^{\frac{I}{2}}\right)^2.$$

That is, it pays on grounds of efficiency for agents to be *sorted* into firms that are hierarchies of polyarchies, each consisting of homogeneous agents.

Analogously, if two firms are polyarchies of hierarchies, then expected output with two identical firms, both of which consists of $\frac{I}{2}$ agents of each type, is greater than the sum of expected outputs of two firms, each of which is consists of identical agents:

$$2\left(1-\left(1-p_1^{\frac{I}{2}}\right)\left(1-p_2^{\frac{I}{2}}\right)\right) \geq \left(1-\left(1-p_1^{\frac{I}{2}}\right)^2\right) + \left(1-\left(1-p_2^{\frac{I}{2}}\right)^2\right).$$

Therefore, it pays on grounds of efficiency for the agents to be *mixed*, when organizations are polyarchies of hierarchies and each of the hierarchies being staffed by identical agents.

3.6.4 A View from the Theory of Teams

An application of results from the theory of teams provides intuitive support of our findings. Prat (2002) considers a problem that is similar to section 3.6, that is whether organizations should employ people with similar or different backgrounds, and derives results that rest on the super-modularity or sub-modularity properties of an organization's payoff function. It is straightforward to show, using Prat's definitions [Prat (2002), p. 1193], the following in the case of organizations with two types of agents:

Proposition 9.

For organizations made up of individuals of two types, two-agent hierarchies are supermodular and two-agent polyarchies are submodular in agents' screening probabilities. Similarly, two-level hierarchies of polyarchies with I_1, I_2, members and screening performance (p_1, p_2) are supermodular, and polyarchies of hierachies, with polyarchies made up of $\frac{I}{2}$ groups of two-level hierarchies, are submodular, in agents' screening probabilities.

Our results above that hierarchies (polyarchies) require that different agents be sorted (mixed) are broadly consistent with Prat's findings that organizations whose payoff functions are supermodular (submodular) in agents' actions should (should not) involve identical actions, if feasible. Furthermore, it turns out that hierarchies of polyarchies are supermodular, and polyarchies of hierarchies are submodular, in agents' screening probabilities. Our results that agents should be *sorted*, if organizations are hierarchies of polyarchies (due to supermodularity), and *mixed*, if organizations are polyarchies of hierarchies (due to submodularity) are again broadly consistent with his findings. Thus by appealing to Prat's results, we conjecture that endogenizing individuals' screening performance would not alter our findings regarding sorting and mixing. Finally, invoking Prat's Proposition 3, we conjecture that if agents' screening outcomes are perfectly co-dependent, supermodular organizational structures are favored, while if they are are independent, submodular organizational structures are favored.

4 A Cognitive Model of Screening

The fact that cross-compositions of pure Sah–Stiglitz architectures are characterized by sigmoid functions of individual performance was shown to be critical in our theory of organizational architecture. We now turn to a model where individual screening behavior that aim at expressing cognitive aspects of agents' decision making and exhibit sigmoid properties. The term *cognitive* model is used somewhat loosely (though in the same sense as in psychology) in order to describe information processing associated with an individual's psychological functions as they pertain to project evaluation. Each individual's decisions combine intrinsic assessment of projects with recommendations from her co-workers. It allows us to study the performance of general organizational architectures where as before organizational design maps different ways to aggregate information. It also allows us to examine a particularly important role of feedback from the apex to the bottom of an organization.

We model an individual's screening performance in terms of the magnitude of a scoring or screening index. For individual i, we define the parameters and components of the scoring index

$$h_i^j + J_i \Omega_{\nu(i)} + \epsilon_{ij}$$

as follows:

1. h_i^j, $j = g, b$, is individual i's own assessment of a project, a "score" that is partially informative provided that $h_i^g > h_i^b$ (good projects receive higher scores than bad projects);

2. $\Omega_{\nu(i)}$ stands for an aggregator of recommendations from $\nu(i)$, the set of agents who report to i (i's subordinates), with $\Omega_{\nu(i)} = 1$ (-1), indicating acceptance (rejection);[16]

3. J_i denotes a positive relative weight individual i assigns to the recommendations from subordinates; and

4. ϵ_{ij} is a random factor, to be discussed shortly.

Let us assume, without loss of generality, that for individual i that the contribution of project quality to the performance index satisfies that $h_i^b < 0 < h_i^g$. Individual i accepts a project of quality j if the value of the scoring index $h_i^j + J_i\Omega_{\nu(i)} + \epsilon_{ij}$ exceeds a certain threshold. By subsuming this threshold in the definitions of (h_i^g, h_i^b), we say that agent i accepts a project, $\omega_i = 1$, if $h_i^j + J_i\Omega_{\nu(i)} + \epsilon_{ij} > 0$; otherwise, the agent rejects it, $\omega_i = -1$. Therefore, agent i's decision, conditional on project quality and on the aggregation of subordinates' decisions, is described by a binary indicator, defined in terms of the distribution function of the ϵ_{ij}'s. Agent i acts in terms of $h_i^j + J_i\Omega_{\nu(i)} + \epsilon_{ij}$. An agent's cognitive ability, as far as screening

[16] Whereas the assumption of the particular numerical values $(1, -1)$ does seem innocuous at this point, and has been used by others in the recent Condorcet Jury Theorem literature (Koriyama and Szentes (2007), those values are quite critical in simplifying some of the analytics, as we see further below.

is concerned, is reflected in the fact that given a recommendation from her subordinates, an agent is more likely to accept a good project than a bad project.

Assuming for convenience that the random factor in the scoring index is extreme-value distributed, the probability that agent i approves of a project of quality j, $j = g, b$, given the recommendation from subordinates, $\Omega_{\nu(i)}$, is expressed in terms of the logistic integral:

$$\text{Prob}\left(\omega_i = 1)|\Omega_{\nu(i)}, j\right) = \frac{\exp\left[\beta\left(h_i^j + J_i\Omega_{\nu(i)}\right)\right]}{1 + \exp\left[\beta\left(h_i^j + J_i\Omega_{\nu(i)}\right)\right]},\ j = g, b. \tag{10}$$

Here β, a positive parameter, represents the sensitivity of the evaluation with respect to the random factors. The case of $\beta = 0$ implies purely random choice: the two outcomes are equally likely. The higher is β, the smaller is the variance of the random factor associated with the evaluation, and the outcome is deterministic if $\beta \to \infty$. Since this specification is arbitrary, given that $h_i^g > h_i^b$, h_i^b can be low enough so as to yield that a bad project is less likely to be accepted than a good project:

$$\text{Prob}\left(\omega_i = 1)|\Omega_{\nu(i)}, g\right) > \text{Prob}\left(\omega_i = 1)|\Omega_{\nu(i)}, g\right).$$

From the results reported in the paper so far, it should not come as a surprise that a critical feature that the cognitive model brings to the analytics is that an agent's screening performance according to Equ. (10), being defined in terms of the cumulative distribution function of the random component in the valuation, is itself a sigmoid function of the recommendations of her subordinates. This formulation is somewhat reminiscent of the properties of informational cascades with large numbers of agents [Bikhchandani, *et al.* (1992)]. However, our model is more general than theirs. In their model, an informational cascade occurs if an individual's action does not depend on his private information signal [*ibid.*, p. 1000]. This amounts to a restriction on the choice model

4.1 General Organizational Architectures

Here we introduce a framework that may be used to examine organizations which are characterized by interdependence among agents which may not presuppose any of the Sah–Stiglitz architectures. We represent organizational architectures by means of directed graphs over the set of agents representing its members, $\mathcal{I}$. Such a general graph-theoretic description of organizations may accommodate easily different notions of authority. E.g., let sets $\mathcal{I}_k,\ k = 1, \ldots, K,\ \mathcal{I} = \bigcup_k \mathcal{I}_k$ be mutually exclusive partitions of the set of all agents, $\mathcal{I}$, that represent the sets of agents at the different hierarchy levels; $\mathcal{I}_1$ is a singleton, the apex of the hierarchy and $\mathcal{I}_K$ is the bottom; $\nu(i)$ is the set of agents who report to i with $|\nu(i)|$ being their number, or i's in-degree; and $\mu(i)$ is the agent whom i reports to, agent i's boss, that is each node has out-degree equal to 1, $|\mu(i)| = 1$. So, if $i' \in \nu(i)$, then $\mu(i) = i'$. And, if $i \in \mathcal{I}_k$, then $\nu(i) \subset \mathcal{I}_{k+1}$. Graph direction represents authority, flowing from superior, indicated by lower numerical values of k, to subordinate, indicated by higher numerical values of k. Essentially, organizations are represented by trees.

Let $\widetilde{\omega}$ denote the $I-$vector containing the decisions of all agents $i = 1, \ldots, I$. Agent i receives recommendations from her subordinates $\nu(i)$, which are stacked as a $|\nu(i)-$vector denoted by $\widetilde{\omega}_{\nu(i)}$. Adapting the cognitive model, we have that agent i makes a decision, taking into consideration her own evaluation and inputs from her subordinates $\Omega_{\nu(i)}$, construed as a general nonlinear function that aggregates inputs from her subordinates $\widetilde{\omega}_{\nu(i)},\ \Omega_{\nu(i)} = \Omega(\widetilde{\omega}_{\nu(i)})$. Decisions by an agent's subordinates are assumed to be independent, conditional on information that they may possess in common and on whether a project is good or bad. Noteworthy special cases are when an agent is sensitive to the average of the recommendations from her subordinates, to unanimity among her subordinates (which corresponds to the Sah–Stiglitz hierarchy), and to whether or not the number of positive recommendations exceeds the number of negative ones by at least a given number. The latter may be interpreted as a concept of decision making by a modified majority; see Sah and Stiglitz (1988).

The unconditional screening probabilities by agent i of a randomly chosen project, $\text{Prob}(\omega_i)$, may be written in terms of the recommendations of agent i's subordinates by working with (10) and the following:

$$\text{Prob}(\omega_i) = \sum_{j \in \{g,b\}, \Omega_{\nu(i)} \in \{1,-1\}} \text{Prob}\left(\omega_i | \Omega_{\nu(i)}, j\right) \text{Prob}\left(\Omega_{\nu(i)} | j\right) \text{Prob}\left(j\right).$$

It is convenient to define

$$\mathbf{\Psi}_{(j)} = \left[\begin{array}{cc} \text{Prob}\left(\omega_i = 1|\Omega_{\nu(i)} = 1, j\right) & \text{Prob}\left(\omega_i = 1|\Omega_{\nu(i)} = -1, j\right) \\ \text{Prob}\left(\omega_i = -1|\Omega_{\nu(i)} = 1, j\right) & \text{Prob}\left(\omega_i = -1|\Omega_{\nu(i)} = -1, j\right) \end{array} \right]. \tag{11}$$

The first (second) row of $\mathbf{\Psi}_{(j)}$, in (11), gives the probabilities for an agent to accept, respectively, to reject, a project of quality j, conditional on receiving a positive (negative) recommendation from her subordinates. From now, to reduce visual cluttering, we simplify notation as follows:

$$\text{P}_{i,\omega_i|\Omega_{\nu(i)},j} = \text{Prob}\left(\omega_i|\Omega_{\nu(i)}1, j\right), \ \text{P}_{i,\omega_i|j} = \text{Prob}\left(\omega_i|j\right), \tag{12}$$

That is, $\text{P}_{i,1|1,j}$, $\text{P}_{i,1|j}$, denote:

$$\text{P}_{i,1|1,j} = \text{Prob}\left(\omega_i = 1|\Omega_{\nu(i)} = 1, j\right), \ \text{P}_{i,1|j} = \text{Prob}\left(\omega_i = 1|j\right).$$

These steps are summarized in the following proposition.

<u>Proposition 10.</u>

The conditional decision probabilities by agent i, $Prob(\omega_i|j)$, as functions of the decisions of her subordinates, $\nu(i)$, $Prob\left(\Omega_{\nu(i)}|j\right)$, $j = g, b$, are given by:

$$\left[\begin{array}{c} \text{P}_{i,1|j} \\ \text{P}_{i,-1|j} \end{array} \right] = \mathbf{\Psi}_{(j)} \left[\begin{array}{c} \text{P}_{\nu(i),1|j} \\ \text{P}_{\nu(i),-1|j} \end{array} \right], \tag{13}$$

and the unconditional ones, under the simplifying notation $P_{i,\omega_i} = Prob\left(\omega_i\right)$, by

$$\left[\begin{array}{c} \text{P}_{i,1} \\ \text{P}_{i,-1} \end{array} \right] = \alpha \mathbf{\Psi}_{(g)} \left[\begin{array}{c} \text{P}_{\nu(i),1|g} \\ \text{P}_{\nu(i),-1|g} \end{array} \right] + (1-\alpha)\mathbf{\Psi}_{(b)} \left[\begin{array}{c} \text{P}_{\nu(i),1|b} \\ \text{P}_{\nu(i),-1|b} \end{array} \right]. \tag{14}$$

We need to specify the aggregator rule $\Omega_{\nu(i)}$, as a function of the set of the recommendations from agent i's subordinates, $\omega_{\nu(i)}$, in order to be able to compute the decision probabilities of the entire organization. Since the organization is defined by means of a directed graph, the probabilities $\text{Prob}\left(\omega_i|j\right)$ may themselves be written by applying (13) recursively in terms of the probabilities describing the decisions of those reporting to each agent i, $\Omega_{\nu(i)}$. We note that (13) may be adapted to apply when screening takes time (and leads in turn to interesting dynamics), as well as when the stochastic dependence occurs contemporaneously and is purely spatial. The tools of Markov random fields may be used in such a case to analyze organizational equilibrium.

Additional results may be obtained for both pure Sah–Stiglitz architectures, when the number of agents is large. The limits for the probabilities of the two possible outcomes for large organizations may be obtained easily. For example, for the hierarchy case, it readily follows by imposing in (13) that the two probability vectors are equal to one another, that $\text{Prob}\,(\omega_i = 1|j) = \text{Prob}\left(\Omega_{\nu(i)} = 1|j\right) = \pi_j$. Then, $(\pi_j, 1-\pi_j)$ is given by the invariant distribution of $\Psi_{(j)}, j = g, b$, which always exists and is positive, as the eigenvector corresponding to the unit eigenvalue of $\Psi_{(j)}$, a stochastic matrix. Solving yields:

$$\pi_j = \frac{\mathrm{P}_{i,1|-1,j}}{\mathrm{P}_{i,1|-1,j} + \mathrm{P}_{i,-1|1,j}}, j = g, b.$$

The organization's decision is given by: $\alpha\pi_g + (1-\alpha)\pi_b$. By expressing this solution in terms of the definition of the conditional choice probabilities in (10), we may write conditions that the parameters of the cognitive model must satisfy in order for the probabilities that a good project be accepted and that a bad project be rejected, $\pi_g, 1-\pi_b$, respectively, to be acceptable in the limit of very large organizations.

We conclude, therefore, that for a large number of agents, the pattern of interdependence among agents, as represented by the stochastic matrix $\Psi_{(j)}$, determines entirely the decisions of the organization. Of course, Ψ^j encapsulates the fundamentals of the cognitive model. We note that in the context of the basic approach of the paper, organizations whose architectures have the graph topology of trees in an intuitive sense also combine the properties of pure hierarchies and polyarchies, as defined in section 2.2 above. That is, a bit like cross-composition, that is motivated by a desire explore complex organizational architectures that by combining pure architectures may enhance their desirable properties and weaken their undesirable ones and is examined in section 3.1, agents who report to the same superior act as in parallel (they act as polyarchies), whereas agents belonging to the same line of authority act as in series (they act as hierarchies). Our cognitive model embedded in general organizational architectures rests critically on how recommendations of subordinates are aggregated at each level. We explore next the importance of the aggregation rule.

4.2 Averaging the Opinions of Subordinates

We assume, for simplicity, that each agent is sensitive to the *average* recommendation of her subordinates, with positive and negative recommendations assuming numerical values $1, -1$, respectively. That is, the expected recommendation of i's subordinates, conditional on project quality j, is $\mathcal{E}\{\Omega_{\nu(i)}|j\} = \frac{1}{|\nu(i)|}\sum_{i'\in\nu(i)} \mathcal{E}\left\{\omega_{i'}|j\right\}$. When the number of subordinates, $|\nu(i)|$ is large, then by the law of large numbers the actual magnitude of $|\nu(i)|$ does not matter [17] and

$$\mathcal{E}\{\Omega_{\nu(i)}|j\} = \mathcal{E}\left\{\omega_{i'}|j\right\},\ i' \in \nu(i).$$

It then follows, for both $j = g, b$, that:

$$\mathcal{E}\left\{\omega_i|j\right\} = \tanh\left(\beta h_i^j + \beta J_i \mathcal{E}\left\{\omega_{i'}|j\right\}\right),\ i' \in \nu(i),\ \forall i \in \mathcal{I}_k,\ k = 1, \ldots, K-1, \tag{15}$$

$$\mathcal{E}\left\{\omega_i|j\right\} = \tanh(\beta h_i^j),\ \forall i \in \mathcal{I}_K, \tag{16}$$

where the hyperbolic tangent function, $\tanh(z) \equiv \frac{e^z - e^{-z}}{e^z + e^{-z}}$, allows us to write concisely the mean of discrete-valued random variables that assume values $1, -1$, and have probabilities given by (10). In the simplest possible case when each level of the hierarchy contains a single agent, $\mathcal{I}_k = 1$ with $K = I$, and all agents are identical, the acceptance probability by agent i, conditional on quality j, is given by (15), with $i+1$ in place of i', and so on down the hierarchy.

The expected profit is given by (2),

$$R_g \cdot \alpha \cdot \mathrm{P}_{\ 1,1|g} + R_b \cdot (1-\alpha) \cdot \mathrm{P}_{1,1|b} - Kw. \tag{17}$$

Given wages, then the optimal number of hierarchy levels is defined in terms of the trade-off between the advantage of additional workers in improving the organization's screening

[17] The case when an agent is sensitive to the actual sample mean of the responses of her subordinates is of particular interest because it models the case of modified majority. In fact, the probability distribution for the sample mean may expressed in terms of the probability distributions for the number of favorable (or of unfavorable) recommendations. Specifically, let Y_i denote the sample mean: $Y_i = \sum_{i\in\nu(i)} \omega_i$. Then, using the fact that favorable recommendations are coded by 1 and unfavorable ones by -1, implies that the number of favorable recommendations is equal to $\frac{Y_i + |\nu(i)|}{2}$. The probability distribution for this random variable may be expressed in terms of the probability of successes in $|\nu(i)|$ Bernoulli trials.

performance and their contribution to cost.[18] Proposition 11 summarizes our results.

Proposition 11.

For a hierarchy of K agents with a single agent in each level, we have the following:

Part A. *The acceptance probability for the entire organization is given by the $K-1$ iterate of the* $\tanh(\cdot)$ *as follows:*

$$\text{Prob}\{\omega_1 = 1|j\} = \frac{1}{2}\left[1 + \mathcal{T}^{K-1}anh\left(\beta h^j\right)\right], \tag{18}$$

where $\mathcal{T}^{K-1}anh$ denotes the $K-1$ order iterate of the hyperbolic tangent function $\tanh(\cdot)$. *They are defined, as $K \to \infty$, by the upper-most fixed point $\overline{\varpi}$ of $\varpi = \tanh\left[\beta h^g + \beta J\varpi\right]$, and by the lower-most fixed point $\underline{\varpi}$ of $\varpi = \tanh\left[\beta h^b + \beta J\varpi\right]$, for good and bad projects, respectively.*

Part B. *Given a wage rate w, there exists a unique number of hierarchy levels, and of agents, K for which expected profit, given by (17), is maximized, provided that:*

$$\frac{1}{2}\left\{R_g\alpha\left[\tanh\left(\tanh\left(\beta h^g\right)\right) - \tanh\left(\beta h^g\right)\right] + R_g(1-\alpha)\left[\tanh\left(\tanh\left(\beta h^b\right)\right) - \tanh\left(\beta h^b\right)\right]\right\} > w.$$

Part C. *If parameters h^g, h^b are sensitive to the wage rate and satisfy*

$$\frac{\partial h^g}{\partial w} > 0, \ \frac{\partial^2 h^g}{\partial w^2} < 0; \ \frac{\partial h^b}{\partial w} < 0, \ \frac{\partial^2 h^g}{\partial w^2} < 0,$$

then there exists a unique wage rate for which expected profit, given by (17), is maximized.

Fixed points $(\overline{\varpi}, \underline{\varpi})$ can always be computed in an iterative fashion from the above formulas. The optimal size of the firm is well defined. They may not be affected by increasing the number of agents, but may be affected by varying the parameters (h^g, h^b). Thus, $\overline{\varpi}$ increases and $\underline{\varpi}$ decreases, causing an improvement in the organizational screening performance at the cost of increased wage cost.

[18] We may endogenize the wage rate by making the parameters of the cognitive model dependent on it. For example, we may assume that increasing the wage rate makes workers more attentive and increases their propensity to tell good, by increasing h^g, and bad, by decreasing h^b, projects.

4.3 Organizations with Top-to-Bottom Feedback

We turn next to examining organizational screening performance when there is feedback from higher levels to lower levels of the hierarchy. We contemplate situations where the principal, agent $i = 1$, directs lower level staff to projects after her own preliminary assessment of their quality. Such feedback from agents at higher levels of an organization to those at lower levels are reminiscent of a model of social interactions in organizations like that of DeMarzo, Vayanos and Zwiebel (2003). [19] The present model's feature that a principal influences the evaluation process at the lowest level of the hierarchy resembles a "closure" in the listening structure in *ibid.*. Such a feature models the role of an organization's principal in influencing an organization's agenda by inputting to entry level decisions. A key consequence of this assumption is, as we see shortly, that the principal's input to entry level decisions influences all decisions and by cascading back up to her own level may render organizational equilibrium *independent* of the number of agents.

Proposition 12. *For a hierarchy of K agents with a single agent in each level, with feedback from agent 1 to agent K, we have the following:*

Part A. *The expected decisions of agent 1 and of agent K, conditional on project quality j, satisfy*

$$E\left\{\omega_1|j\right\} = \mathcal{T}^{K-1}anh\left(\beta h^j + \beta J E\left\{\omega_K|j\right\}\right), \tag{19}$$

$$E\left\{\omega_K|j\right\} = \tanh\left(\beta h^j + \beta J_{1K} E\left\{\omega_1|j\right\}\right). \tag{20}$$

A unique pair of values (ϖ_K^j, ϖ_1^j) satisfying (19)–(20) always exists, $j = g, b$.

[19] In DeMarzo, Vayanos and Zwiebel (2003), individuals update their own beliefs by weighting the beliefs of the other agents they "listen to" in terms the respective relative precisions. In their dynamic model in the long run beliefs converge to a set of consensus beliefs which depend only on the spectral properties of the matrix that describes the listening structure. In the special case of hierarchical organizations possessing a "standard listening structure," that is where every agent except the chief has a single boss and each agent listens to her subordinates and to all of those located higher up than herself, the relative beliefs over the frequency distribution that characterizes consensus beliefs of an agent and a subordinate depend only on the numbers of her subordinates and on her level in the hierarchy [*ibid.*, Theorem 10].

Part B. *If $J = J_{1K}$, in which case the system of fixed point equations (19)–(20) is symmetric, then organizational screening performance is independent of the number of agents.*

The screening probabilities for good, respectively, bad projects are given by the same formulas as above and are concave increasing, respectively, convex decreasing functions of the number of hierarchy levels. The screening probabilities are given by $\frac{1}{2}\left[1 + \tanh\left(\beta h^g + \beta J \overline{\varpi}^g\right)\right]$, for good projects, and $\frac{1}{2}\left[1 + \tanh\left(\beta h^b + \beta J \underline{\varpi}^b\right)\right]$, for bad projects, where $\overline{\varpi}^g$ is the upper fixed point of $\varpi = \tanh(\beta h^g + \beta J \varpi)$, and $\underline{\varpi}^b$ is the lower fixed point of $\varpi = \tanh(\beta h^b + \beta J \varpi)$.

The symmetry with which all agents are treated and the fact that the model rests only the average decision are clearly responsible for the resulting independence from the number of agents. This implies, in turn, that the number of hierarchy levels, and thus the number of agents as well, is indeterminate.

This result allows us to conclude that the cognitive model of individual behavior allows us to avoid the complexity of the Von Neumann–Moore–Shannon solution and ensures the existence of a finite-sized organizations that satisfy arbitrary performance standards. Still, it is interesting that top-to-bottom feedback improves screening performance over any finite-sized hierarchy, by allowing an organization to attain immediately the best possible values for the likelihood of rejecting good projects and of accepting bad ones as equilibrium outcomes. Still, this comes at the cost of ultimate constraints on organizational performance.

The results of this example and of the earlier one on averaging over the opinions of subordinates both demonstrate an important property of the cognitive model: organizational screening performance is a sigmoid function of $h^j, j = g, b$, each agent's contribution to her own evaluation of a typical project. With the cognitive model, adding agents improves over individual screening performance with pure Sah–Stiglitz architectures, such as a hierarchy, even *without* mixing (that is, cross-composing) of architectures. Organizational architecture then may be optimized in order to attain other objectives of the firm.

5 Conclusion

This paper models organizations whose members screen projects. It examines how architectural features of organizations affect organizational screening performance and shows that for an organization with several agents to be able to improve upon individual performance, the organization's screening function must be sigmoid in individual performance. Even without a cognitive model of individual performance, organizations with mixed Sah–Stiglitz architectures, such as hierarchies made up of components that are polyarchies and polyarchies made up of components that are hierarchies, do give rise to such functions. This property is in turn critical for the determination of optimal firm size and hierarchy levels, and the endogenous determination of individual screening performance. Hierarchies and hierarchies of polyarchies with heterogeneous agents are supermodular, and polyarchies and polyarchies of hierarchies are submodular, in individual screening performance. These properties allows us to link with the theory of teams and to show that with heterogeneous agents hierarchies (polyarchies) require that different agents be sorted (mixed). We also show that agents should be sorted, if organizations are hierarchies of polyarchies, and mixed, if organizations are polyarchies of hierarchies. These results are consistent with the theory of teams and provide concrete ways to construct organizations with supermodular or submodular performance. The paper also introduces a cognitive model that allows for individuals' own screening efforts to combine with influence from others, subordinates and superiors. The paper examines the implications of such interactions for the limits to organizational performance. The paper also examines general properties of committee decision making and shows the superiority of committees with respect to composition, that is when each member of organization is replaced by a replica of the entire organization. The paper links with the modern Condorcet Jury Theorem literature, but such a link has not been made formally by the literature and deserves further attention in future research. It would also be interesting to introduce discretion along the decision making process, like the sort of discretion prosecutors have in judicial proceedings or journal editors and other organizations (such as research funding agencies) in evaluating projects.

APPENDIX: Proofs

Proposition 1. Proof. Condition $\frac{\partial}{\partial p}\mathcal{P}(p^*) > 1$ is established as follows. From Theorem 1, Moore and Shannon (1956), P. 198, we have:

$$\frac{\partial}{\partial p}\mathcal{P} > \frac{(1-\mathcal{P})\mathcal{P}}{(1-p)p}.$$

Applying it at the fixed point p^* yields

$$\frac{\partial}{\partial p}\mathcal{P}|p^* > \frac{(1-\mathcal{P}(p^*))\mathcal{P}(p^*)}{(1-p^*)p^*} = \frac{(1-p^*)p^*}{(1-p^*)p^*} = 1.$$

Clearly, after one round of composition, we have that:

$$\mathcal{P}(p_b) < p_b < p^* < p_g < \mathcal{P}(p_g). \tag{21}$$

In other words, the organization improves screening in both directions. Q.E.D.

Proposition 2. Proof.

Part A follows by differentiating (3) with respect to p.

$$\frac{\partial \mathcal{P}_B(p;k+1,K)}{\partial p} = \frac{p^k(1-p)^{K-k-1}}{\int_0^1 t^k(1-t)^{K-k-1}dt};\ \frac{\partial^2 \mathcal{P}_B(p;k+1,K)}{\partial p^2} = \frac{p^{k-1}(1-p)^{K-k-2}}{\int_0^1 t^k(1-t)^{K-k-1}dt}(k-(K-1)p).$$

The approval probability assumes the value of 0, at $p=0$ and of 1, at $p=1$. Its derivative with respect to p is zero at both $p=0,1$ and positive throughout $(0,1)$. Its second derivative with respect to p is positive for $p < \frac{k}{K-1}$ and negative $p > \frac{k}{K-1}$. Therefore, it has a unique fixed point in $(0,1)$, at which in view of the properties just established, the graph of $\mathcal{P}_B(\cdot)$ must cross with a slope greater than 1. Committee approval requires distinguishing between K's being even, $K=2N$, in which case $k-(K-1)p$ above becomes $N-(2N-1)p$; or odd, $K=2N+1$, in which case $k-(K-1)p$ becomes $N-(2N)p$, with the minimum number of members constituting a majority being $N+1$ in either case. The inflection point is $\tilde{p} = \frac{k}{K-1}$, which for the case of even number of agents is equal to $\frac{N}{2N-1} \approx \frac{1}{2}$. The slope of the approval probability at its inflection point is:

$$\frac{\left(\frac{1}{2}\right)^{2N-2}}{\int_0^1 t^k(1-t)^{K-k-1}dt}.$$

The key question here is how this slope varies when K increases. Employing Stirling's approximation for the beta function in the denominator of (3) and above, $\mathcal{B} = (N+1, N)$, yields:

$$\frac{p^{2N-2}}{\int_0^1 t^N (1-t)^{N-1} dt} \approx \frac{4}{\sqrt{\pi}} N^{\frac{1}{2}} \frac{\left(\frac{1}{2}\right)^{2N+\frac{1}{2}}}{\frac{(N+1)^{N+\frac{1}{2}} N^N}{(2N+1)^{2N+\frac{1}{2}}}}.$$

The slope varies with N according to $N^{\frac{1}{2}}$. Therefore, as N increases, the slope increases, too.

Part B is handled conveniently by employing the normal approximation to the binomial distribution. The probability of k successes in K independent Bernoulli trials is given by a normal distribution with mean Kp and variance $Kp(1-p)$. The probability that a $K-$ member committee approves of a project is given by:

$$\mathcal{P}_B\left(p; \frac{K}{2}+1, K\right) \approx 1 - \Phi\left(\frac{K^{\frac{1}{2}}\left(\frac{1}{2}-p\right)}{p^{\frac{1}{2}}(1-p)^{\frac{1}{2}}}\right), \tag{22}$$

where $\Phi(\cdot)$ denotes the standardized normal distribution function. This approximation is more accurate the larger is K and the closer is p to $\frac{1}{2}$. It is straightforward to derive that

$$\frac{\partial}{\partial K}\mathcal{P}_B\left(p; \frac{K}{2}+1, K\right) \approx -\phi\left(\frac{K^{\frac{1}{2}}\left(\frac{1}{2}-p\right)}{p^{\frac{1}{2}}(1-p)^{\frac{1}{2}}}\right)\frac{1}{2}\frac{K^{-\frac{1}{2}}\left(\frac{1}{2}-p\right)}{p^{\frac{1}{2}}(1-p)^{\frac{1}{2}}},$$

$$\frac{\partial^2}{\partial K^2}\mathcal{P}_B\left(p; \frac{K}{2}+1, K\right) \approx \frac{1}{4}\phi\left(\frac{K^{\frac{1}{2}}\left(\frac{1}{2}-p\right)}{p^{\frac{1}{2}}(1-p)^{\frac{1}{2}}}\right)\frac{K^{-\frac{3}{2}}\left(\frac{1}{2}-p\right)}{p^{\frac{1}{2}}(1-p)^{\frac{1}{2}}}\left[1+\frac{K^{\frac{1}{2}}\left(\frac{1}{2}-p\right)}{p^{\frac{1}{2}}(1-p)^{\frac{1}{2}}}\right],$$

where $\phi(\cdot)$ denotes the standardized normal density function. The claims above readily follow.

Q.E.D.

Proposition 3. Proof.

This is a restatement of Theorem 2, Moore and Shannon (1956), p. 202–204. The generality of the result comes from the fact that both $\mathcal{P}(p)$ and $F_q(p)$ are defined in terms of Boolean functions of the same number of variables. The general form of $\mathcal{P}(p)$ is a polynomial: $\mathcal{P}(p) = \sum_{m=0}^{n} A_m p^m (1-p)^{n-m}$, where n is the number of agents. Here, A_m is the number of ways one can select a subset of m agents, such that if these agents approve of a project, and the remaining agents disapprove, the project is rejected by the organization. The quorum function may also be written in a polynomial form which is a restricted one relative to that of $\mathcal{P}(p)$, but in such a way that comparisons are possible. This allows Moore and Shannon to prove their Theorem 2.

Q.E.D.

Proposition 5. Proof.

The first derivative of the screening function (6) with respect to p is

$$\frac{\partial}{\partial p}\mathcal{P}_{HP} = I(1-p)^{\frac{I}{K}-1}\left(1-(1-p)^{\frac{I}{K}}\right)^{K-1},$$

is strictly positive in $(0,1)$, and assumes the value of 0 for $p = 0, 1$. Its second derivative,

$$\frac{\partial^2}{\partial p^2}\mathcal{P}_{HP} = \frac{I}{K}(1-p)^{\frac{I}{K}-2}\left(1-(1-p)^{\frac{I}{K}}\right)^{K-2}\left[K(I-1)(1-p)^{\frac{I}{K}} - (I-K)\right],$$

assumes the value of 0 for $p = 0, 1$, and is positive for small values of p and negative for high values of p. It thus admits a unique inflection point, $\tilde{p}_{HP} = 1 - \left(\frac{I-K}{K(I-1)}\right)^{\frac{K}{I}}$, $\tilde{p}_{HP} \in (0,1)$, provided that $I > K > 1$. Since $\mathcal{P}_{HP}(0) = 0$, $\mathcal{P}_{HP}(1) = 1$, the screening function is indeed sigmoid over $[0,1]$.

Therefore, $\mathcal{P}_{HP}$ has a fixed point in $(0,1)$, which satisfies $p^*_{HP} = \left(1-(1-p^*_{HP})^{\frac{I}{K}}\right)^K$. It readily follows in view of its properties that the screening function intersects the 45^o line from below at its unique fixed point p^*_{HP}. The fixed point p^*_{HP} is a decreasing function of I.

Q.E.D.

Proposition 6. Proof.

The screening function (7) is sigmoid with respect to individual screening performance follows from the fact that its value is equal to 0, at $p = 0$, and is equal to 1, at $p = 1$, and that its first derivative of the screening function with respect to p,

$$\frac{\partial}{\partial p}\mathcal{P}_{PH} = I\left(1-p^K\right)^{\frac{I}{K}-1}p^{K-1}$$

is always positive in $(0,1)$, but becomes equal to 0 for $p = 0, 1$. Its second derivative

$$\frac{\partial^2}{\partial p^2}\mathcal{P}_{PH} = I(1-p^K)^{\frac{I}{K}-2}p^{K-2}\left[K-1-(I-1)p^K\right],$$

is positive for small values of p and negative for high values of p. Therefore, it admits an unique inflection point: $\tilde{p}_{PH} = \left(\frac{K-1}{I-1}\right)^{\frac{1}{K}}$, in $(0,1)$.

Therefore, $\mathcal{P}_{PH}$ has a unique fixed point in $(0,1)$, which satisfies

$$\left(1-(p^*_{PH})^K\right)^{\frac{I}{K}} = 1 - p^*_{PH}.$$

The properties of (7) with respect to I, K follow readily. To establish that the screening function intersects the 45^o line from below at its unique fixed point p^*_{PH}, we need to determine the magnitude of $\frac{\partial}{\partial p}\mathcal{P}_{PH}(p^*_{PH})$.

Q.E.D.

Proposition 7. Proof.

The optimal wage rate must satisfy the first- and second-order conditions:

$$\alpha \cdot R_g \cdot \mathcal{P}'\left[P_g(w)\right] P'_g + (1-\alpha) \cdot R_b \cdot \mathcal{P}'\left[P_b(w)\right] P'_b = I, \tag{23}$$

$$\alpha \cdot R_g \cdot \left(\mathcal{P}''\left[P_g(w)\right](P'_g)^2 + \mathcal{P}'\left[P_g(w)\right] P''_g\right) + (1-\alpha) \cdot R_b \cdot \left(\mathcal{P}''\left[P_b(w)\right](P'_b)^2 + \mathcal{P}'\left[P_b(w)\right] P''_b\right) < 0. \tag{24}$$

Note that a sufficiently high value of the wage rate makes the first term within the first set of parentheses in (24) negative, because of the concavity of $\mathcal{P}[P_g(w)]$ for high values of w. It also makes the first term within the second set of parentheses positive, because of the convexity of $\mathcal{P}[P_b(w)]$ for high values of w, making it more likely that the entire second term be negative. As the wage rate increases, each agent becomes more adept in recognizing good projects, thus increasing $P_g(w)$, and also more adept in recognizing bad projects, thus decreasing $P_b(w)$.

Q.E.D.

Proposition 8. Proof.

While we may deal with the full generality of the problem of maximizing expected profit, as in Proposition 6, we simplify by considering the special case of $R_b = 0$. Then the problem is simply to maximize with respect to $(w_{k_1}, w_{k_2}; I_{k_1}, I_{k_2})$ the screening probability given by (4), with $p_k = P_g(w_k)$, minus the total wage cost, given by (8). The first-order conditions with respect to (w_{k_1}, w_{k_2}) are:

$$P_g'(w_{k_1})\,(1-P(w_{k_1}))^{I_{k_1}-1}\left(1-(1-P(w_{k_2}))^{I_{k_2}}\right)=1; \tag{25}$$

$$P_g'(w_{k_2})\left(1-(1-P(w_{k_1}))^{I_{k_1}}\right)(1-P(w_{k_2}))^{I_{k_2}-1}=1. \tag{26}$$

By combining these two conditions, we can see claim (a) in inequality form follows. Since the screening probability is increasing concave in (w_{k_1}, w_{k_2}), a unique solution exists. It also follows that if $I_{k_1} = I_{k_2}$, then claim (a) also holds in equality form.

The first-order conditions with respect to (I_{k_1}, I_{k_2}) are:

$$-\ell n\,(1-P(w_{k_1}))\,(1-P(w_{k_1}))^{I_{k_1}}\left(1-(1-P(w_{k_2}))^{I_{k_2}}\right)=w_{k_1}; \tag{27}$$

$$-\ell n\,(1-P(w_{k_2}))\left(1-(1-P(w_{k_1}))^{I_{k_1}}\right)(1-P(w_{k_2}))^{I_{k_2}}=w_{k_2}. \tag{28}$$

Since the screening function is increasing concave in (I_{k_1}, I_{k_2}), a unique solution exists.

By substituting from (25), (26), respectively into (27), (28), respectively, we get:

$$\frac{\partial \ell n\,(1-P(w_{k_1}))}{\partial w_{k_1}}\frac{1}{\ell n\,(1-P(w_{k_1}))}=\frac{1}{w_{k_1}},$$

and the corresponding one in terms $P(w_{k_2})$. It follows that (w_{k_1}, w_{k_2}) are fully determined from the respective equations. Therefore, if all wage rates and numbers of agents are variable, then symmetry suggests that the same number of agents are used at all levels and all agents are remunerated at the same wage rate. This proves claim (b).

Q.E.D.

Proposition 9. Proof.

Let there be two types of agents with screening abilities $(\hat{p}_1, \hat{p}_2), (\check{p}_1, \check{p}_2)$. Assuming in turn that $(\hat{p}_1 > \check{p}_1,\ \hat{p}_2 > \check{p}_2)$, and that $(\hat{p}_1 > \check{p}_1,\ \hat{p}_2 < \check{p}_2)$, we can show by means elementary comparisons that:

$$\hat{p}_1\hat{p}_2+\check{p}_1\check{p}_2 \le \max\{\hat{p}_1,\check{p}_1\}+\max\{\hat{p}_2,\check{p}_2\}+\min\{\hat{p}_1,\check{p}_1\}+\min\{\hat{p}_2,\check{p}_2\},$$

so that the two-agent hierarchy is supermodular in (p_1, p_2);

$$1-(1-\hat{p}_1)(1-\hat{p}_2)+1-(1-\check{p}_1)(1-\check{p}_2) \ge 1-(1-\max\{\hat{p}_1,\check{p}_1\})+1-(1-\min\{\hat{p}_1,\check{p}_1\})\,(1-\min\{\hat{p}_1,\check{p}_1\}),$$

For hierarchies of polyarchies, we can show that:

$$\left(1-(1-\hat{p}_1)^{I_1}\right)\left(1-(1-\hat{p}_2)^{I_2}\right)+\left(1-(1-\check{p}_1)^{I_1}\right)\left(1-(1-\check{p}_2)^{I_2}\right)$$

$$\leq 1-(1-\min\{\hat{p}_1,\check{p}_1\})^{I_1}+1-(\max\{\hat{p}_2,\check{p}_2\})^{I_2},$$

which confirms supermodularity.

Finally, for polyarchies of hierarchies, can show that:

$$1-(1-\hat{p}_1\hat{p}_2)^{\frac{I}{2}}+1-(1-\check{p}_1\check{p}_2)^{\frac{I}{2}}$$

$$\geq 1-(1-\min\{\hat{p}_1\check{p}_1\}\min\{\hat{p}_2\check{p}_2\})^{\frac{I}{2}}+1-(1-\max\{\hat{p}_1\check{p}_1\}\max\{\hat{p}_2\check{p}_2\})^{\frac{I}{2}},$$

which confirms submodularity.

Q.E.D.

Proposition 11. Proof.

Part A.

It is straightforward to establish analytically that adding levels in the hierarchy improves organizational screening performance. That is, Prob$\{\omega_1 = 1|g\}$ is an increasing concave function of K, and Prob$\{\omega_1 = 1|b\}$ is decreasing convex function of K, the number of hierarchies (and agents). Therefore, $R_b < 0$ suffices for expected revenue to be concave in the number of agents and therefore for the optimal number of agents to be well defined, given w.

It is easiest to do this graphically. Note that:

$$\text{Prob}\{\omega_1^j = -1|g\} = \frac{1}{2}\left[1+\tanh\left(\beta h^g+\beta J E(\omega_2)\right)\right], \tag{29}$$

$$\text{Prob}\{\omega_1^j = 1|b\} = \frac{1}{2}\left[1+\tanh\left(\beta h^b+\beta J E(\omega_2)\right)\right]. \tag{30}$$

Starting with the agent at level K, the bottom of the hierarchy, the probability of rejecting a good project is given by $\frac{1}{2}\left[1-\tanh\left(\beta h^g\right)\right]$, and the probability of accepting a bad project is given by $\frac{1}{2}\left[1+\tanh\left(\beta h^b\right)\right]$. These probabilities are equal $\frac{1}{2}|G_1G^1|$ and $\frac{1}{2}|B_1B^1|$, respectively, in Figure 2.

The rejection probability for a good project after adding a second agent is given by:

$$\text{Prob}\{\omega_1 = 1|g\} = \frac{1}{2}\left[1 - \tanh\left(\beta h^g + \beta J\left(\tanh\left(\beta h^g\right)\right)\right)\right].$$

It then follows that as the number of levels K increases, the acceptance probability for a good project increases. It tends asymptotically to 1 minus the value of the upper most point of $\varpi = \tanh\left[\beta h^g + \beta J\varpi\right]$, when K tends to infinity. This fixed point, $\overline{\varpi}$, always exists and is stable if $h^g > 0$ [Ioannides (2006), Proposition 1].

We work similarly for the acceptance probability for a bad project. Referring to Figure 2, the map of $\tanh\left(\beta h^b + \beta J\varpi\right)$ goes through the point with horizontal coordinate equal to $-\beta h^b$ (recall, $h^b < 0$, for simplicity, though it is not necessary). We note the value of $\tanh\left(\beta h^b\right)$, is given from the map of $\tanh(\omega)$ by $-|b_1 B_1|$. The probability of accepting a bad project is $\frac{1}{2}|B_1 B^1| = \frac{1}{2}\left[1 + \tanh\left(\beta h^b\right)\right]$. Then $\tanh\left(\beta h^b + \beta J \tanh\left(\beta h^b\right)\right)$ may be read off the $\tanh\left[\beta h^b + \beta J\underline{\varpi}\right]$ map and marked on the horizontal axis. Clearly, by construction, $|B_2 B^2| < |B_1 B^1|$, and therefore the probability of accepting a bad project decreases when a second individual evaluates it. As K increases tending to infinity, the acceptance probability for a bad project tends asymptotically to $\frac{1}{2}[1 + \underline{\varpi}]$, where the lower fixed point satisfies $\underline{\varpi} = \tanh\left[\beta h^b + \beta J\underline{\varpi}\right]$, and is negative. In view of these explanations, expression (18) readily follows from (15) and (16).

Part B.

We conclude that since the upper and lower fixed points of the respective screening functions are well defined, they provide the lower bounds for the screening probabilities. As more agents are added, the acceptance probability of good projects is increasing and for bad projects is decreasing by decreasing increments that tend to zero. That makes expected revenue concave in K. The condition under Part B ensures that the largest marginal revenue from an additional hierarchy level exceeds w, the marginal cost. The former tends to 0 as K tends to infinity. This ensures existence of a unique organization size K^*.

Part C.

Differentiating expected profit with respect to w involves the expressions for the acceptance probabilities (29–30) and thus the properties of the tanh function. These functions are concave in w in the appropriate range. The F.O.C. for a maximum involves the derivative of tanh, $\tanh(z)' \equiv 4\left(e^{z}+e^{-z}\right)^{-2}$, which assumes a maximum of $\beta\frac{\partial h^j}{\partial w},\ j=g,b.$

Q.E.D.

<u>Proposition 12. Proof.</u>

Part A. Proposition 2 in Ioannides (2006) again applies with minimal modification and ensures that there always exist upper solutions $\overline{\varpi}^g_K, \overline{\varpi}^g_1)$ of Equ. (19) – (20) for $j = g$. Similarly, lower solutions $(\underline{\varpi}^b_K, \underline{\varpi}^b_1)$, of Equ. (19) – (20) for $j = b$. always exist. Therefore, the optimal size of the organization may be established along the lines of Proposition 8.

Part B. Suppose that agents at level K expect agent 1's average decision to be $\hat{\omega}^j_1$, conditional on a project of quality j. Then the average decision for agents at level K is $\mathcal{E}\left\{\omega_K\right\} = \tanh\left(\beta h^j + \beta J \mathcal{E}\left\{\hat{\omega}^j_1\right\}\right)$. It follows then that the expectation of agent 1's decision satisfies

$$\mathcal{E}\left\{\omega^j_1\right\} = \mathcal{T}^{K-1}anh\left(\beta h^j + \beta J \hat{\omega}^j_1\right),$$

where $\mathcal{T}^{K-1}anh[\cdot]$ denotes the $K-$th order iterate of the function $\tanh\left(\beta h^j + \beta J\varpi\right)$. Therefore, the principal's decision as an organizational outcome at equilibrium satisfies $\mathcal{E}\left\{\omega^j_1\right\} = \hat{\omega}^j_1$, and is given a fixed point of the function $\mathcal{T}\mathrm{anh}^K$. The fixed points of this function coincide with the fixed points of the function $\varpi = \tanh(\beta h^j + \beta J\varpi)$. Ioannides (2006) shows that there may either one or three fixed points, depending upon parameter values. Therefore, we conclude that due to "closure" in the listening structure (made possible by the top level's providing feedback directly to the lowest level of the hierarchy), the screening probabilities at equilibrium are independent of the number of levels of the hierarchy.

Q.E.D.

REFERENCES

Austen-Smith, David, and Jeffrey S. Banks (1996), "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *The American Political Science Review*, 90, 1, 34–45.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *The Journal of Political Economy*, 100 (5), 992–1026.

Boland, Philip J. (1989), "Majority Systems and the Condorcet Jury Theorem," *The Statistician*, 38, 181–189.

Le Marquis de Condorcet, Marie Jean Antoine Nicolas Caritat (1785; 1994), *Essai sur l' application de l' analyse à la probabilité des décisions rendues à pluralité des voix,* Les Archives de la Revolution Française, Pergamon Press, 1994.

DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel (2003), "Persuasion Bias, Social Influence and Uni-Dimensional Opinions," *Quarterly Journal of Economics,* 118, 909–968.

Fioretti, Guido, and Bauke Visser (2002), "A Cognitive Approach to Organizational Complexity," University of Siena and Erasmus University working paper, October 2002.

Haykin, Simon (1999) *Neural Networks,* Prentice Hall, Upper Saddle River, NJ.

Ioannides, Yannis M. (1987), "On The Architecture of Complex Organizations," *Economics Letters*, 25, 201–206.

Ioannides, Yannis M. (2006), "Topologies of Social Interactions," *Economic Theory* 28, 559-584.

Koriyama, Yukio, and Balázs Szentes, (2007), "A Resurrection of the Condorcet Jury Theorem," working paper, Department of Economics, University of Chicago, November 13.

Kremer, Michael, and Eric Maskin (1996), "Wage Inequality and Segregation by Skill," NBER working paper No. 5718, August.

MacLeod, W. Bentley (2002), "Complexity, Bounded Rationality and Heuristic Search," *Contributions to Economic Analysis and Policy, The B. E. Journals in Economic Analysis and Policy,* 1, 1.

Mount, Kenneth R., and Stanley Reiter (1998), "A Modular Network Model of Bounded Rationality," Chapter 8, 306–340, in Majumdar, Mukul, ed., *Organizations with Incomplete Information: Essays in Economic Analysis, Tribute to Roy Radner*, Cambridge university Press, New York.

Mount, Kenneth R., and Stanley Reiter (2002), *Computation and Complexity in Economic Behavior and Organization* , Cambridge university Press, Cambridge.

Moore, E. F., and C. E. Shannon (1956), "Reliable Circuits Using Less Reliable Relays," *Journal of the Franklin Institute,* 262, Part I, 191–208, Part II, 281–297.

Papadimitriou, Christos H. (1994), *Computational Complexity,* Addison Wesley Longman, Reading, MA.

Prat, Andrea (2002), "Should a Team Be Homogeneous?" *European Economic Review,* 46, 1187–1207.

Reiter, Stanley (1996), "Coordination and the Structure of Firms," Northwestern University, Discussion paper No. 1121 R, September.

Sah, Raaj K. (1991) "Fallibility in Human Organizations and Political Systems," *Journal of Economic Perspectives,* 5, 2, 67–88.

Sah, Raaj K., and Joseph Stiglitz (1985), "Human Fallibility and Economic Organization," *American Economic Association Papers and Proceedings,* 75, 2, 292–297.

Sah, Raaj K., and Joseph Stiglitz (1986), "The Architecture of Economic Systems: Hierarchies and Polyarchies," *American Economic Review,* 76, 4, 716–727.

Sah, Raaj K., and Joseph Stiglitz (1988), "Committees, Hierarchies and Polyarchies," *The Economic Journal,* 98, 391, 451–470.

Simon, Herbert A. (1960), "The Architecture of Complexity," *Proceedings of the American Philosophical Society,* 106, 6, 467–482.

Visser, Bauke (2000), "Organizational Communication Structure and Performance, *Journal of Economic Behavior and Organization,* 42, 2, 231–252.

Visser, Bauke (2001a), " Complexity, Robustness and Performance: Trade–Offs in Organizational Design," Erasmus University working paper, September.

Visser, Bauke (2001b), " Exploring Organizational Complexity," Erasmus University working paper, September.

von Neumann, John (1956), "Probabilistic Logics and The Synthesis of Reliable Organisms from Unreliable Components," *Automata Studies, Annals of Mathematical Studies,* 34, 43–98.

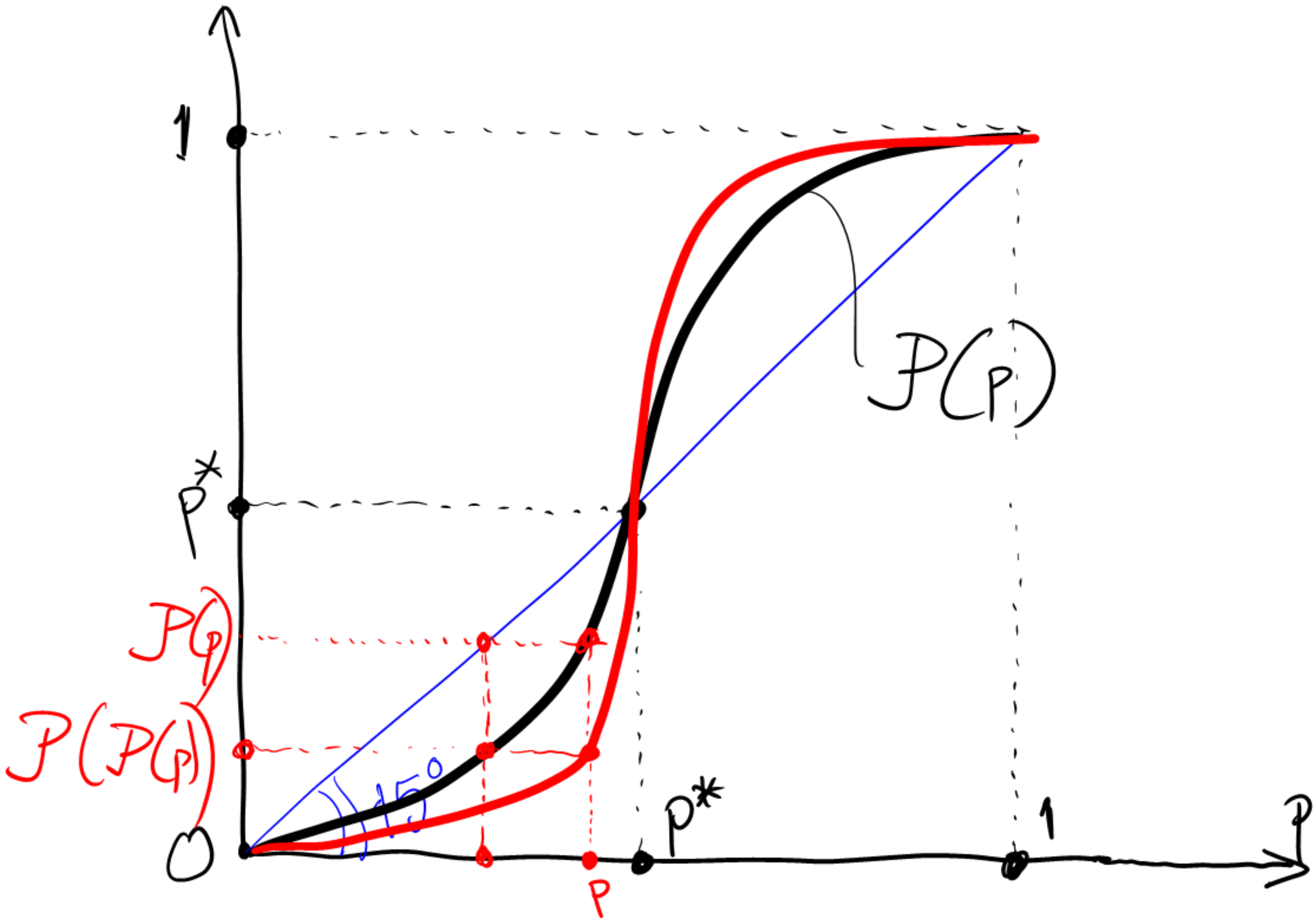

FIGURE 1