F Brahm and J Poblete
The evolution of productive organizations
Article

Brahm, F and Poblete, J

(2021)

The evolution of productive organizations.

Nature Human Behaviour, 2021 (5). pp. 39-48. ISSN 2397-3374

DOI: https://doi.org/10.1038/s41562-020-00957-x

Nature Research (part of Springer Nature)
https://www.nature.com/articles/s41562-020-00957-x

The Evolution of Productive Organizations

Francisco Brahm[1*] and Joaquin Poblete[2,3]

[1] London Business School, London, UK (orcid id: https://orcid.org/0000-0002-2529-6392)

[2] School of Management, Pontificia Universidad Catolica de Chile, Santiago, Chile (orcid id: https://orcid.org/0000-0002-2546-4261)

[3] Instituto Sistemas Complejos de Ingeniería, Santiago, Chile

* Corresponding author: Francisco Brahm (fbrahm@london.edu)

1 Abstract

Organizations devoted to the production of goods and services, such as guilds, partnerships and modern corporations, have dominated the economic landscape in our species' history. We develop an explanation for their evolution drawing from cultural evolution theory. A basic tenet of this theory is that social learning, under certain conditions, allows for the diffusion of innovations in society, and therefore, the accumulation of culture. Our model shows that these organizations provide such conditions by possessing two characteristics, both prevalent in real world organizations: exclusivity of membership and more effective social learning within their boundaries. The model and its extensions parsimoniously explains the cooperative nature of the social learning advantage, organizational specialization, organizational rigidity and the locus of innovation. We find supportive evidence for our predictions using a sample of pre-modern societies drawn from the Ethnographic Atlas. Understanding the nature of these organizations informs the debate about their role in society.

2 Introduction

Organizations, defined as a stable and interacting collection of individuals with a common and specific goal[1], have played a crucial role throughout human history. One type of organization, which we define as "productive organization" (PO), focus on producing and delivering the goods and services that satisfy the needs of human populations (e.g., food, shelter, clothes, tools). Whether it is the societas in Roman times[2], the guilds in medieval times[3], the partnerships in early Renaissance[4], or the modern corporation[5], POs have played a crucial role in our species' success. Not surprisingly, Herbert Simon noted that our "market economies" are in reality "organizational economies"[6].

An often neglected corollary is that the theories about the nature of these organizations have an important impact on public policies about markets and organizations, as well as how these organizations are managed. Consider, for example, the consequential debate around the role of corporations in society[7,8]. However, while extant theories explain how POs work (i.e., their inner workings), we lack a clear explanation for their evolution. This is not a minor gap, as understanding POs' origin is a requisite to fully understand, and thus to improve the capacity to harness, manage or regulate this dominant institution.

In this paper, we propose such an explanation drawing from Cultural Evolution Theory[9-14]. Cultural evolution studies the transmission and inheritance of culture, defined as information –beliefs, norms, knowledge, skills, and techniques– acquired from other individuals via social learning (e.g., imitation, teaching). A central insight of this theory is that social learning, by way of diffusing innovations in a society, produces their gradual accumulation over time. This process of cultural accumulation has allowed humans to adapt and conquer every environment in the globe. We propose that POs evolved because they enhance the capacity of social learning to generate cumulative culture. This contributes to the nascent discussion on the importance of studying organizational evolution and its impact on cumulative culture[15-18].

Our model and its extensions can illuminate several aspects of POs that have eluded an integrated and parsimonious explanation so far. First, we use culture as the main explanatory lever, away from the focus on incentives and governance in the existing (economic) theories of POs. These theories propose that POs' role is to avoid the myriad of transaction hazards involved in market exchange[19-24]. For each type of hazard, POs provide a distinct solution. For example, when assets are specific to a transaction and therefore hold-up is likely, POs resolve ex-post conflict using authority[22]. While empirical evidence supports this approach[25], recent evidence shows that POs are, to a large extent, carriers and transmitters of culture[26-30]. Our theory informs the latter: it proposes that two specific and ever-present characteristics of POs –restricted access and improved social learning–, favor cultural accumulation in societies. Thus, while current theories assume a pre-existing cultural pool from which transactions emerge, we endogenize culture and POs.

Second, the importance of cooperation for POs' nature is a point that has been frequently made[6,18,31-34] but scarcely formalized[17,35]. Our theory proposes that cooperation is a cultural quality of POs that evolves via group selection, instead of being enforced upon self-interested agents via incentives and other governance devices. For this, we use the idea of teaching as a cooperative act that enhances the effectiveness of social learning [14,36,37].

Third, we lack an understanding of the evolutionary mechanisms that selected POs. Current theories focus on how POs work, that is, they provide a 'proximate explanation' by detailing the mechanisms of governance and incentive provision[20,22,24,38,39]. The spread of POs is then explained by invoking agents' causal understanding whom adopt POs to increase their expected utility. However, causal understanding doesn't need to be present to produce complex cultural phenomena[13,40,54] and thus, POs might not be adopted via complex foresight and calculations. Instead, as with other traits that societies carry and transmit over time, entrepreneurs may simply adopt POs as the inherited default at their disposal. Therefore, if one assumes limited causal understanding in humans[6,13,40], asking how, when and why POs originated and spread is pertinent. This is the type of explanation we provide in this paper, known as 'ultimate explanation' in biology[39,41,42]. Just as biologists study the function of a trait for reproduction and survival, we unpack the POs' function for cumulative culture, the key enabler of our species' success[40,13,14]. Our ultimate explanation of POs complements evolutionary approaches in economics[43] –that focus on change but not origin– and, by unpacking the role of exclusivity and enhanced social learning, it provides a clue that can guide empirical work in cultural evolution, be that historical[2,29,44,45], cultural-phylogenetic[46] or archeological[47]. In this paper, we provide a step in that direction by testing the model using a sample of pre-modern societies drawn from the Ethnographic Atlas.

Fourth, POs display a high level of specialization, that is, they execute a limited set of tasks and activities producing a narrow output. An example is specialization across guilds: weavers, painters, masons, bookbinders, leatherworkers, candle makers, among many others. Current theories of POs do not explain this specialization; instead, they require an additional theoretical mechanism: specialized POs can benefit from comparative advantage –generated endogenously by repetition[48,49] or by exogenous resource endowment[50]– by trading with other specialized POs. In this paper, we propose an alternative account for the rise of specialization, namely, it maximizes the social learning advantage of POs, and therefore the benefits they bring to society. This can explain the prevalence of specialized guild-like organizations prevalent in ancient forager societies which had limited access to trade[51].

3 Results

3.1 Formal model

We use a workhorse model in the cultural evolution literature[52,53]. (Mathematical proofs are presented in the supplementary information.) In every period, a continuum of long-lived agents adopt a technology that confers fitness, but whose value is subject to changes in the environment. There are N environmental states. In each period, the state may change with probability p. For every state, there is a unique technology that provides a positive level of fitness, while the remaining technologies provides no fitness at all. Because we model fitness in terms of difference between variants, by a normalization, we can assume without loss of generality that the fitness of a technology is 0 unless it is tuned to match the state (in which case fitness is 1). Agents adopt a technology by using one of two behavioral strategies. An Individual Learner studies and understands her environment and is able to develop in each period a new technology tuned to the current state. This strategy has a cost C, which is bounded between 0 and 1. The second alternative is to learn socially. A Social Learner looks at what some other randomly chosen member of the population did in the previous period and simply copies its technology, incurring a cost $c < C$. This strategy is less costly because the agent does not need to understand the underlying state but rather simply copies what others do. In order for social learning to survive, we assume $(C - c) > p$. Within a period, events occur in the following order: all agents observe the state of the nature; social learners pick a random agent and copy what she did in the past; individual learners execute their learning; and fitness levels are realized.

Let r_I be the share of individual learners and r_S the share of social learners in the population, and let q be the percentage of people with a tuned technology in the population. The variable q is governed by the difference equation $q(t+1) = r_I + q(t) \cdot (1-p) \cdot r_S$. For any given pair of shares r_I and r_S, the expected ratio of tuned agents is given by $q^e(r_I, r_S) = \frac{r_I}{1-(1-p)r_S}$, the steady state of the difference equation. The fitness of an individual learner is $f_I = (1 - C) > 0$ because she is always tuned to the state bearing the cost C. Across the paper we assume weak selection, and thus, the fitness of a social learner will approximate in the long run to $f_S = (1-p)q^e - c$. Social learners, given that they copy their behavior from others, sacrifice tuning if the state of the world changes or if inadvertently they copy from an untuned member. The fitness of social learners is increasing in the share of individual learners because the chances of copying from tuned individuals grow when there are more individual learners.

We assume that at the end of every period, after fitness are realized, a small proportion of agents adopt the strategy with higher fitness level. This type of evolutionary dynamics is known as a quasi birth and death process and converges to evolutionarily stable strategies (ESSs). (We could also assume that agents are short-lived and their reproduction rate depends on their fitness level; in either case, the equilibrium concept is ESS, and our results are the same.) Given weak selection, a population with shares (r_I, r_S) plays an ESS if a small group of invaders using any different share of these strategies achieves a strictly lower average fitness. Consistent with prior literature, we find that there exists a unique equilibrium in which the share of individual learners is given by $r_I = \frac{p\,[1-(C-c)]}{(1-p)(C-c)}$, and the fitness of both types is $1 - C$. Intuitively, if there are mostly individual learners, social learners would benefit from a high q with a fraction of the cost, and therefore they would expand; in contrast, if there are mostly social learners, they would be worse off because, even if they have a low cost, they have a high risk of becoming technologically outdated due to low q in the population.

Observe that because both behavioral types in equilibrium achieve the same fitness $1 - C$, this implies that society as a whole does not benefit from the existence of social learning. The same average level of fitness can be achieved with individual learning only. The fact that social learning is selected but does not affect the fitness of the population is known as Roger's paradox[52]. This result has demonstrated to be robust to different specifications and assumptions, and has generated an important literature exploring the conditions that makes social learning a source of adaptive cumulative culture[14,53,54]. In what follows, we show that adding a PO in the society solves this paradox in a way that, we argue, is fundamentally different from other solutions proposed in the literature because it does not require that social learning enhance the fitness of individual learners[53,54]. We introduce first a single PO in order to facilitate the understanding of the mechanisms at play; then, we introduce multiple POs to enrich the model with cooperation and group selection ideas.

Productive organizations. Two characteristics define a PO. First, access to the PO is limited. A fixed fraction λ of agents is located inside the organization. This means that even though additional members might want to be a part of the PO, membership is exclusive, limited by λ. This speaks to a fundamental characteristic of POs: "organizations, as a condition of their existence, must maintain boundaries that separate them from their environments. In the absence of distinguishable boundaries, there can be no organization as we understand them"[1].

Exclusiveness is a persistent characteristic of POs across history: Neolithic sodalities, Roman societas, medieval guilds, Renaissance partnerships and modern day firms all limit the possibility of becoming a member. Exclusiveness appears to be present from the first records of POs. For example, evidence shows that the production daggers in Scandinavian society in the late Neolithic was "consciously organized to keep the recipes of the technology exclusive to certain segments of the society"[55]. A review of sodalities –probably the first non-kin goal-oriented organizations[56]– indicates sodalities have "no common characteristics beyond the fact that they all excluded non-members"[57,58].

Second, if a member of the PO copies another member it bears cost $\tilde{c}$ which is lower than c; if she copies from outside, she bears a cost c. This means that within PO the effort to transmit information is lower and/or that the fidelity of the information transmission is higher. Information transmission such as teaching or mentoring is a cooperative act[14,36,37]. To achieve cooperation within POs we first invoke some exogenous determinants such as: λ might create population structure which generates spatial-selection/assortment of cooperatives types[59-60]; membership to a PO might trigger deep-seated tribal instincts that boosts group-identity and help among fellow group members[17,18,61-63]; POs can add formal governance structures (e.g., authority, rules) as cooperation enforcement devices[22,64]. We also endogenize cooperation in an extension of the model by using the idea of group selection among multiple competing POs[65,17] (see below and the SI for details).

Empirical evidence strongly supports more effective social learning within POs. The management and economics literature has documented that learning from coworkers of the organization is more effective than learning from the outside[27,66] and the guilds and partnerships improve knowledge transmission between its members[29,30]. In archeology, evidence from three centuries of amphorae production in workshops in the Roman Empire shows that the variability of amphorae between workshops is consistent with a process of high-fidelity social learning within workshops (i.e., master to disciples) instead of horizontal transmission or mobility between workshops[47]. In forager societies, a review of the evidence on learning shows that "craft expertise -- the kind of skill sets that forager lives depend on-- is fine-tuned at a generation and reliably transmitted across generations by this mode of organized human learning environments" and "the parallels with the formal, institutionalized system of apprentice guilds could hardly be clearer"[51].

Our definition of PO maps well to the idea of clubs in economics[67-69]. Just like a club, a PO entails the private provision of a public good that can be made exclusive (lower cost of social learning), but where each PO member imposes an externality on the rest (higher exposure to a change in the environment).

As before, let r_I be the share of individual learners outside the PO and $\tilde{r_I}$ be the share of individual learners inside the PO. In the same fashion define r_S, $\tilde{r_S}$, q, $\tilde{q}$, q^e and $\tilde{q}^e$. The dynamics of q and $\tilde{q}$ are ruled by the following difference equations:

$$q(t) = r_I + r_S\,(1-p)[\lambda\,\tilde{q}(t-1) + (1-\lambda)\,q(t-1)] \tag{1}$$

$$\tilde{q}(t) = \tilde{r_I} + \tilde{r_S}\,(1-p)[\lambda\,\tilde{q}(t-1) + (1-\lambda)\,q(t-1)] \tag{2}$$

Social learners imitate randomly from the whole population (below we study different assumptions for social learning). From (1) and (2), the steady state values for the tuned population are the following:

$$q^e = \frac{r_I + (1-p)\,\lambda\,(r_S - \tilde{r_S})}{1-(1-p)(\lambda\,\tilde{r_S} + (1-\lambda)\,r_S)} \tag{3}$$

$$\tilde{q}^e = \frac{\tilde{r_I} + (1-p)(1-\lambda)(\tilde{r_S} - r_S)}{1-(1-p)(\lambda\,\tilde{r_S} + (1-\lambda)\,r_S)} \tag{4}$$

If $\lambda = 0$, we return to the steady state of the basic model. These equations clarify the intuition that the PO, in addition to creating the benefit of cheaper social learning, also generate a reduction of the average tuning of the population if $\widetilde{r_S} > r_S$.

To understand the equilibrium and shares of different strategies, let's introduce their fitness. The fitness of an individual learner is the same outside or inside the PO, $\widetilde{f_I} = f_I = (1 - C) > 0$. The fitness of a social learner outside the PO is,

$$f_S = (1-p)[(1-\lambda) \cdot q^e + \lambda \cdot \tilde{q}^e] - c \tag{5}$$

Given $\widetilde{c} < c$, the social learner inside the PO enjoys a higher fitness equal to,

$$\widetilde{f_S} = (1-p)[(1-\lambda) \cdot q^e + \lambda \cdot \tilde{q}^e] - [(1-\lambda)c + \lambda\widetilde{c}] = f_S + \lambda(c - \widetilde{c}) \tag{6}$$

We assume that outside the PO, a replicator dynamic modifies the share of the two strategies according to their relative payoff; the same occurs inside the PO, in every period a small percentage of agents change strategy depending on their payoff. If the PO is small, the evolutionary dynamic generates an equilibrium where: i) the expected fitness of both behavioral strategies outside the PO equalize and both strategies have a share, ii) given that $\widetilde{f_S} > \widetilde{f_I}$, only social learners populate the PO (see SI for proof) (see panel D of figure 1). The latter implies a trade-off: while a PO populated only by social learners takes full advantage of cheaper social learning, it decreases the average level of tuning in the population (panel B of figure 1).

However, this lower tuning requires a large PO to outweigh the benefits of cheaper social learning. Using i), ii) and equations (3) to (5), we find the share of social learner outside the PO to be equal to $r_I = \frac{[1-(C-c)]\,p}{(C-c)(1-p)} \cdot \frac{1}{(1-\lambda)}$. Thus, the expansion of the size of a PO increases the share of individual learners outside the PO (panel C of figure 1). (At low levels of λ, the total share of individual learners in the population is the same as in the basic model, namely $\lambda \cdot 0 + (1-\lambda) \cdot r_I = \frac{[1-(C-c)]\,p}{(C-c)(1-p)}$.) The intuition for this is that, with a PO populated only by social learners, social learners outside the PO learn more from social learners, and thus, are more exposed to environmental change. This decreases their fitness and favors the expansion of individual learners. This rise in r_I compensates the decrease in tuning of POs from having only social learners, generating the linear increase in PO fitness depicted in panel A of figure 1. However, if λ continues increasing a point is reached where there are only individual learners outside the PO and thus, the fitness of the PO is maximal (see panels A and C of figure 1). After this point, if λ increases, the impact of lower tuning generated by PO comes to dominate the social learning benefits it generates, and thus the fitness of the PO decreases. If the PO becomes too large, then it ceases to benefit the population at all.

From this discussion it is easy to see that, provided that there is a positive measure of social learners outside the PO (which happens if the PO is not too big), there exists a unique equilibrium in which the average fitness of society is greater than that of individual learners. The result is stated formally in the next proposition:

Proposition 1. If λ is sufficiently small, the existence of the PO increases the average fitness of the population.

The basic intuition for this result is that the PO limits the negative externality of social learners. In the basic model, social learners reproduce and grow, lowering the average fitness of the population q until the fitness decreases to the level of individual learners $(1 - C)$. With the introduction of a PO of limited size, this expansion is put to a halt before all the benefits of cheaper social learning are diluted away.

Some implications. We highlight three implications of our model: locus of innovation, organizational inertia, and rules for optimal PO size.

First, given that a fitness-enhancing PO is populated only by social learners, the locus of innovation in our model, is outside the PO. This is consistent with the common view that, over history, radical innovations – the discovery and application of basic laws of nature (i.e., individual learning)– have tended to happen outside POs, for example, by inventors, entrepreneurs, and individual scientists. Recent evidence shows that this is also the case for modern firms: most manufacturing firms, when introducing a "new to the world" product (i.e., a truly new product or technology), they source the invention from external and independent inventors, such as scientists, labs/inventors, small start-ups and users[70]. With this, we are not saying that within POs there is no innovation; we simply say that individual learning –by relating primarily to the concept of radical innovations– is more likely to occur outside the PO. We execute three extensions of the model that create a role for individual learning and innovation inside the POs (available in the SI). First, by adding selective learning we find that POs do engage in individual learning, but at a lower proportion than outside POs (see the robustness section below for more details). Second, POs thrive in generating incremental innovations –not radical innovation– as the technology gets marginally improved every time it is transmitted and copied between its members (this matches the idea of "guided variation" in cultural evolution[10]). There is sizeable literature on organizational learning documenting that these incremental improvements can be substantive[27,66,71], enlarging the positive impact that POs can have on society. Third, in figures S8 and S9 of the SI we show that a model where the PO also reduces individual learning cost, but less than the reduction in social learning costs, yields a PO that is first populated by individual learners when small, and then, as it grows, it becomes fully populated by social learners. This is consistent with the usual life cycle of firms: small startups founded and managed by innovators eventually become standardized big corporations, geared to the replication of tried-and-tested processes and products.

Second, given that the PO is populated only by social learners they bear a relatively larger risk of environmental change, and thus, have a harder time adapting when a change occurs (see figure S1 in SI). This fact is consistent with widespread evidence on organizational rigidity and inertia in sociological and management research[1,72].

Third, panel A of figure 1 shows that there exists a value λ^* that maximizes the fitness of the members of the PO. This maximum will be attained endogenously if an additional assumption is made in our model. Observe that, regardless of the size of the PO, people will always weakly prefer to belong to a PO. Therefore, to endogenize PO size, it is sufficient to assume that new members will be admitted (or expelled) as long as doing it increases the average fitness of the members of the PO. This is not an implausible assumption, as guilds and partnerships have historically maximized the members' average benefit[73-75]. In the SI, we show, consistently with prior research[75], that the size of the PO that maximizes society's fitness is larger than λ^*.

-------- Figure 1 around here --------

Comparative statics. The model provides interesting comparative statics, which are depicted in figure 2. Given that the fitness contribution of the PO comes from facilitating social learning, it is intuitive to find that increases in p (panel C of figure 2) and increases in $\widetilde{c}$ (panel B of figure 2) generate a decrease in the fitness of POs. Less intuitively, we find that changes in the cost of social learning outside the PO (parameter c) do not translate into a monotonic change in the PO's fitness; instead, a decrease in c decreases fitness in small PO but increases it in a large PO (panel A of figure 2).

------- Figure 2 around here ---------

Robustness to different assumptions for social learning. Social learners in our model only imitate, and they do it randomly. However, in practice, individuals tend not to be stuck on a specific strategy –they are selective– and social learning can be biased[10,11,14,76]. In order to increase the ecological validity of our findings, we explored how three types of social learning biases and selective learning affect our model (see supplementary information for the details).

First, we study PO-biased social learning, that is, social learners inside the PO might preferentially learn from fellow members of the PO to take advantage of the lower cost. We find that proposition 1 (and its implications) holds provided that the extent of the bias is not very large. This is consistent with the organizational literature, where attending to knowledge from the environment –away from an exclusive inward focus– is necessary for organizational success[1,27,77], and with the existence of journeymanship in guilds, where artisans would travel to another city in search of new skills, thereby compensating the inward focus of the guild[3,29].

Second, we study secrecy as a particular case of biased social learning. We define a PO to be secret if social learners outside the PO cannot imitate members of the PO. We find that under secrecy the proposition 1 still holds (and its implications), but it that reduces the fitness impact of POs (see panel D of figure 2). The reason is that the share of individual learners outside the PO no longer increases as λ expands (instead, it reverts back to $r_I = \frac{[1-(C-c)]\,p}{(C-c)(1-p)}$) and social learners inside the secretive PO learn from a less tuned population. This suggest that secrecy, a frequent characteristic of real POs, must bring additional benefits in order to evolve, such as a lower cost $\tilde{c}$, which might come, for example, from galvanizing "us versus them" dynamics. Also, it suggest that secrecy, while it motivates individual inventors, it might have the counteracting effect of reducing aggregate innovation in society; evidence from second world war secrecy in the US provides evidence consistent with this[78].

Third, we study pay-off biased social learning, that is, social learners prefer to imitate individuals that are more successful, or, in our case, have higher fitness[10,79]. We explore a model where a portion ϕ of social learners copies the group with the highest payoff (individual learners, social learners outside the PO or social learners inside the PO), and the remainder copy unbiasedly. The portion ϕ captures the extent of the payoff bias. We find that: i) consistent with the literature[54], payoff bias doesn't solve Roger's paradox in model of a single technology, and ii) the addition of a PO yields the same results: the PO increases the fitness of the society and is populated only by social learners. Further, we find that the fitness of the PO is enhanced by the extent of the payoff bias. This happens because payoff bias helps to overcome the main disadvantage of the PO which is slower adaptation; with payoff bias, tuned technologies flow quicker into the PO.

Fourth, we explore a canonical model of selective learning, that is, the capacity of individuals to switch between individual and social learning depending on the circumstances[11,40]. In this model, individuals obtain a cue from the environment and evolution selects the optimal threshold for this cue, above (below) which individual (social) learning is selected. Thus, if the optimal threshold is larger (smaller), social learning is more (less) prevalent. While selective learning is known to solve Rogers' paradox[54], and thus the PO is not the crucial condition to produce cumulative culture in this model, all the other results remain unchanged: the PO enhances the fitness of the society, the proportion of social learning is larger inside the PO than outside, and the dynamics of PO size are the same: the benefit of cheaper social learning is traded-off with the increasing exposure of an environmental change.

Multiple POs, Cooperation and Group Selection. The theory of cultural evolution has put forward the idea of cultural group selection (or multilevel selection)[10,17,61,62,65,80,81]. Several processes generate stable heterogeneity between groups. Cooperative groups have an advantage against non-cooperative groups, and thus, selection can favor within group cooperation and altruism even if it is costly for the individual. We use this idea to extend our model to study multiple competing POs within a society, as frequently occurs in practice.

First, we use group selection to justify the origin of a lower social learning cost inside POs. In the extended model (see the SI for details), we specify four possible strategies as both individual learners and social learners can be altruistic: by bearing a cost δ they can reduce the social learning cost that the copier/imitator is experiencing from c to $\tilde{c}$ (e.g., teaching and guiding). The parameters δ and $\tilde{c}$ are such that altruism is individually disadvantageous but collectively beneficial. The fitness of a particular strategy within PO *i* depends on its own payoff and the average payoff inside the PO, and the weight on the latter captures the strength of group selection pressure. This extension yields the following proposition:

Proposition 2: If the strength of the group selection is high enough and the cost to benefit ratio of altruism is low enough, then the POs are populated only by altruistic social learners which reduce the cost of social learning and thus, the POs are able to increase the fitness of society.

While one PO in a society may still be beneficial if cooperation is exogenously driven by other mechanisms[59,60], this extension to the model endogenizes the origin of the social learning advantage of POs and clarifies its cooperative nature.

Second, we explored the consequences of group selection. In a complementary extension (see SI for details), we assume that multiple POs have different capacity to promote cooperation and thus to lower its social learning costs. We show that more cooperative POs will attain a larger size in equilibrium. Thus, group selection is not only fundamental for POs' learning advantage but also for the expansion within society of more advantageous POs.

Finally, there is still an additional selection level at play. POs have spread and invaded almost all societies across the globe, and thus, selection at the level of societies might have also played a role[13,81]. A society with exclusive and cooperative POs would be more adaptive and successful; thus, different group selection processes between societies –imitation between societies, migration, differential reproduction rate, or size advantage in warfare– propelled the spread of POs further.

Specialization. Suppose now there are M technologies which are assumed to be ex ante equally productive and that become less useful if nature changes with probability p. The strategies of social and individual learners now specify a superindex $j \in J$ that denotes that an agent's learning occurs in a particular technology j. Individual learners are uniformly distributed among activities and social learning occurs randomly but is restricted to the set of people that executed the same technology j in the previous period. A natural way in which this "clustering" of agents around technologies might occur in real societies is that subsets of social learners follow and group close to specific individual learners, and that the resulting groupings are distant, either spatially or socially (e.g., via clans or casts). The cost of individual and social learning and the fitness functions are the same as the case of one technology.

The results of this model are straightforward and mirror the case with only one technology. Social learners are selected into the population, but the overall fitness of the society does not increase. Given our assumptions, individual and social learners will be distributed evenly across technologies. (This would change trivially if individual learners are not distributed uniformly.) The total share of social learners will decrease with c and p.

We now allow for the existence of one PO of size λ. A social learner with technology j will now bear a cost $\tilde{c}$ when imitating agents using technology j whom are sharing the PO, and cost c (larger than $\tilde{c}$) when learning from agents with technology j located outside the PO. As before, the PO generates advantages in social learning.

To show our main result, define $x^j = r_S^j + r_I^j$ as the share of agents that execute technology j outside the PO and analogously define $\tilde{x}^j = \tilde{r}_S^j + \tilde{r}_I^j$ as the share of people inside the PO that executes technology j. The fitness of a social learner outside the PO is given by

$$f_S^j = (1-p)\left[\frac{\lambda\,\tilde{x}^j}{\lambda\,\tilde{x}^j + (1-\lambda)x^j}\;\tilde{q}^{j^e} + \frac{(1-\lambda)x^j}{\lambda\,\tilde{x}^j + (1-\lambda)x^j}\;q^{j^e}\right] - c \tag{3}$$

The fitness of social learners inside the PO can be written as follows:

$$\tilde{f}_S^j = f_S^j + (c - \tilde{c})\frac{\lambda\,\tilde{x}^j}{\lambda\,\tilde{x}^j + (1-\lambda)x^j} \tag{4}$$

The fitness advantage of social learners of technology j within the PO is increasing in the share of members in the PO that execute the same technology. Therefore, whichever technology takes the lead with a higher $\tilde{x}^j$ in a particular period –for example due to a random shock– will enjoy a higher fitness, and thus, will be replicated at a faster pace after each period; this, in turn, will generate an additional advantage in fitness in the next period and thus, an even quicker replication, and so on, for each subsequent period. Consequently, this positive feedback loop will drive one technology to dominate the PO. (And, as an exclusive PO provides advantage to social learners, only these will populate it.). This is stated formally in the next proposition:

Proposition 3. Given a sufficiently small λ, the PO specializes in a specific technology.

The intuition for this result is that people that can learn from each other in a cheaper and more effective way will slowly tend to group together. The PO provides the conditions for more effective learning and therefore they group in there.

We can also extent this proposition to multiple POs. In the SI we show that society benefits from their presence and that each PO will specialize in a specific technology.

These results suggest that specialization within POs, a trait observed from their earliest historical account[47,51,55], evolved because it favored the capacity of POs to benefit society via improvements in social learning. Our theory does not rely on trade and comparative advantage between units –countries, cities or organizations– as the driving force of specialization of these units[50]. Nor does specialization become constrained by the extent of the market[48]. Even in small societies with no trade and comparative advantage (i.e., all agents bear the same opportunity cost of doing any technology, or in our model, homogeneous costs $\tilde{c}$, c and C), specialization will evolve within POs.

3.2 Empirical analysis

We execute four empirical exercises that, together, provide suggestive evidence that POs played an important role in making social learning adaptive as our proposition 1 predicts. Our empirical analysis is a "proof of concept", that is, a way to verify that the theory we propose can be empirically productive and that having a cultural evolution theory of POs can facilitate the empirical search for its origins.

Baseline analysis

To test the predictions of the model, we use the Ethnographic Atlas (EA)[82] and the Standard Cross Cultural Sample (SCCS) provided by the D-Place dataset[83] (see methods section for the details of the data). The EA provides information about eleven productive activities (or technologies) in the society (e.g., metal working, pottery making). The dataset identifies whether each activity was present in the society and, if so, whether it was "normally performed by many or most adult men, women, or both" or was "largely performed by a small minority who possess specialized skills". We identify the second condition as the addition of a PO to the execution of a specific activity. We measure PO in this way because the type of 'small minority' covered by the EA fit the requirements of our theoretical model (see Methods section for a discussion of this fit).

We computed two variables: the percentage of activities that are present in the society ("% presence") and the percentage of those activities that are executed within a PO ("% within PO") (see Methods section for details). In the dataset, there is missing information about the activities due, for example, to the fact that the ethnography did not study one or more productive activities. Only 263 societies had complete information about the eleven activities. We added several control variables (see the Methods section) which brought the number of societies included in the analysis to 173. In the supplementary information we provide the details on the geographical distribution and descriptive statistics of the final sample.

To test the impact of the presence of activities and PO on the fitness of the individuals in society i, we use the following econometric model:

$$Population_i = b1 + b2 \text{ x } \%Presence_i + b3 \text{ x } \%Presence_i \text{ x } \%withinPO_i + Controls_i + Error_i$$

Population as a dependent variable captures the standard notion of fitness as reproductive success. It also captures the fact that in pre-modern Malthusian economies, progress translated into increases in population and not per-capita wealth[29]. We proxy population by using the "size of local communities", which is a categorical variable with 8 categories (see Methods section). In the panels A, B and C of figure 3, we plot the unconstrained relationships between "size of local communities", "% presence" and "% within PO".

The ordered probit estimates are presented in table 1. Assuming that there are no POs, the results presented in column 1 of table 1 show that moving from 0% to 100% in the presence of activities is associated positively with an increase in the local population but this result is not statistically significant (coefficient b2 = 1.137; t-test$_{114}$ = 1.39, two-tailed; p-value = 0.165; 95% CI = {-0.466; 2.742}). Although this result might seem surprising, it is consistent with Rogers' paradox: culture (i.e., more activities) does not necessarily leads to increased fitness. However, consistent with proposition 1 of the model, in column 2 we find that activities are associated with an increase in the local population only when POs are present in the society. This increase is statistically significant (coefficient b3 = 4.298; t-test$_{113}$ = 3.36, two-tailed; p-value = 0.001; 95% CI = {-1.641; 1.675}) and its magnitude is depicted by the difference between the red and blue lines in the panel D of figure 3.

-------- Insert table 1 and figure 3 around here --------

Endogeneity

The first threat to identification of causality is omitted variables. We executed a test that uses selection on observables to assess the extent to which selection on unobservables would need to be in order to overthrow the results[84]. Following the pre-established criteria[84], the test indicates a low threat of omitted variables (see supplementary information for details).

The second threat to identification is reverse causality. This threat can be present both in the presence of activities and in the use of POs. To counter these two threats we use an instrumental variable technique. In the supplementary information we detail our instruments, their theoretical logic and the empirical results. We find that the results of table 1 do not change; if anything, they marginally increase in size.

Comparative statics

The EA, the SCCS and Kirby et al (83) provide several variables to explore empirically the comparative statics of our model depicted in figure 2. We use one variable that proxy for p, three variables that proxy for social learning costs, and two that proxy for secrecy. We find supporting evidence across them all for our predicted comparative statics (see the supplementary information for details). Here we highlight one variable: uncertainty. To proxy for p, we use the measure provided by the D-PLACE dataset[83] of year-to-year climate unpredictability between 1901 and 1950, the period that has the largest proportion of ethnographies in the EA. Consistent with our model, we find that the impact of POs on population decreases when climate unpredictability is high.

Robustness

In the supplementary information we execute the following robustness checks: i) we use alternative dependent variables such as "population density" or "total population", ii) we add additional controls that account for two alternative explanations for our findings, namely, the presence in trade in societies drives the emergence of specialized POs and POs are the product of political complexity (i.e., complex chiefdoms), iii) we are less restrictive regarding the number of societies included in the analysis, increasing observations well above the baseline sample of 173. Across all of these checks, the results of table 1 remained unchanged.

4 Discussion

In this article, we have developed a theory that explains the evolution of productive organizations (POs). As with any trait that has been selected in a population, a full explanation of the nature of POs requires adding an evolutionary perspective to the mix. We used a cultural evolution model to show that improvements in social learning within POs can favor the hard-to-propel process of cumulative culture. If access to POs is restricted, as is typical across history, then this advantage in knowledge transmission leads to higher fitness of societies and therefore to the gradual selection and invasion of POs. The advantage in social learning within POs is cooperative in nature, stemming from group selection among competing POs. Specialization of POs emerges naturally in our theory, as it maximizes the social learning benefits of POs. The theory applies straightforwardly to pre-modern POs, such as guilds, and other long-standing POs, such as partnerships; as a descendant of these older POs, our theory also informs the origin of modern firms. We contribute to cultural evolution by highlighting the importance of organizations in cumulative culture –and thus the need to study their evolution– and to extant (economic) theories of POs by focusing on culture and explaining the evolutionary origins of POs.

Theoretically, we show that social learning can escape Roger's paradox even if it does not generate a positive externality on individual learning's fitness[53,54]. Our model can parsimoniously rationalize several enduring characteristics of POs: restriction of access; easier social learning within than between organizations; importance of cooperation for the fitness of POs; social learning (or, tradition) is dominant inside POs while individual learning (or, innovation) is dominant outside POs; specialization of POs; and the inertia and low adaptive capacity of POs. Our theory can provide clues for necessary historical, cultural-phylogenetic and archeological work to proceed[44-47,81]. We began doing this ourselves: we successfully test our theory using data from the EA[82].

We can point to several limitations in our paper. First, the empirical test we perform is informative but not definitive. It would be interesting to test the predictions of our model using emerging datasets on ancient guilds[85]. Second, our theory is well suited for guilds and partnerships, where knowledge and technology is transmitted across individuals. However, modern firms combine specialized knowledge to generate complex technologies that require many individuals to produce and thus transmission is not at the individual level. Our model could be extended to study the unique evolutionary origins of modern firms. Third, while we do address how innovation can occur within POs, a more complete formal analysis could be performed, particularly to study their role in incremental innovation.

5 Methods

Data

The EA describes cultural practices for 1291 pre-modern societies, ranging from societies with complex agricultural economies and political systems to small hunter-gatherer groups. The societies are globally distributed, with especially good coverage of Africa and western North America. The SCCS is a subsample of the EA where additional information about societies is provided. We use the SCCS to measure several variables that are needed to test the predictions of the model. These datasets were created by coding the available information about societies that is present in the extensive ethnographic accounts in the anthropology literature. Data collection and analysis were not performed blind to the conditions of the study.

Fit of EA with our theoretical model

Our model requires that PO possess three characteristics in order to benefit society: improved social learning, small size and exclusivity. In the discussion that follows, the improved social learning of the minorities of the EA become evident; therefore, we don't expand on it. Regarding small size, the very definition in the EA specifies a “small minority”. Exclusivity requires more care to be mapped to the EA. The minorities in the EA can be of four types: senior age specialization (i.e., only men or women beyond the prime of their life), junior age specialization (i.e., only boys or girls before the age of puberty), craft specialization (which includes occupational castes where the rights to execute certain activity were inherited), and industrial specialization (i.e., specialization is removed from age or craft specialization and is executed using industrialized techniques). Aggregating across activities, craft specialization covers roughly 85 percent of the cases, industrial specialization accounts for eight percent and senior/junior specialization split the rest. Industrial specialization and senior/junior specialization comply with the exclusivity criteria. In the former, exclusivity is predicated on employment, and in the latter, it is defined by age. To understand exclusivity in craft specialization, we randomly sampled twenty societies from the EA and read their original ethnographic accounts. Roughly, we could identify three types of craft specialization. The first one, and the most common type, are organizations that could be described as "proto-guilds". These organization were similar to medieval guilds, they had experts, sometimes called "masters", and apprentices, which would come together regularly --or seasonally, for example in fishing at high latitudes-- in order to exchange work for teaching and to learn from each other. Apprentices typically needed to prove their capacity in order to fully access the community of experts, so access was not freely granted. Being a master often carried prestige in the society. Not infrequently, the right to execute a particular craft/activity was hereditary (e.g., fishing in the Chekiang society in China), generating occupational castes (or a specialized clan). However, even with heredity, skill was also a pre-requisite to enter the "proto-guild". Therefore, heredity boosted exclusivity. The second type of craft specialization were "workshops". These were small and scattered production units, where one or more skillful specialists, with the help of a handful of workers, would serve the needs of a portion of the society, typically the local town or region (for example, metalworking in the Riffian culture in Northern Africa). The third type of craft specialization is the "attached specialists" where skilled craftsmen were appointed and funded by the rulers of the society (e.g., metalworking in the Inca Empire). Either by the selection of workers or their funding, the second and third types of craft specialization also seem to ensure exclusivity. All considered, even though there is heterogeneity in the "craft specialization" of the EA, the basic idea exclusivity in these organizations seems to hold ground.

Variable measurement

The variable "% presence" is computed as the division of the count of activities that were present in the society over the count of activities for which we had available information. If the ethnographic atlas indicates "missing data", then we would not consider that activity in the denominator. If the ethnographic atlas indicates "the activity is absent or unimportant in the particular society", then would not consider that activity in the numerator.

The difference between these is that in the former, the original ethnography did not provide any indication regarding the activity.

The variable "% within PO" is computed as the division of the count of activities "largely performed by a small minority who possess specialized skills" over the count of activities that were present in the society. The relationship between "% presence" and "% within PO" is positive, with a correlation coefficient of 0.4.

The dependent variable "size of local communities" is a categorical variable with 8 categories. The categories are: 1 is "less than 50 people" 2 is "from 50 to 99 persons", 3 is "from 100 to 199 persons", 4 is "from 200 to 399 persons", 5 is "from 400 to 1,000 persons", 6 is "more than 1,000 in the absence of indigenous urban aggregations", 7 is "one or more indigenous towns of more than 5,000 inhabitants but none more than 50,000", and 8 is "one or more indigenous towns with more than 50,000 inhabitants".

As controls, we added geographical variables (e.g., absolute latitude), resource endowment variables (e.g., mammal richness), intensity of agriculture dummies (e.g., semi-intensive), region dummies, type of settlement dummies (e.g., nomadic) and year of the ethnographic record. See the supplementary information for details on these variables.

6 Data availability

The source data used in this paper is publicly available at https://d-place.org/. The dataset used in this paper is available at https://sites.google.com/site/fcobrahm/.

7 Code availability

The code used to analyze the data is available at https://sites.google.com/site/fcobrahm/.

8 References

1. Scott, R. *Organizations: Rational, Natural and Open Systems*. (Prentice Hall, Fifth Edition, 2003).
2. Hansmann, H., Kraakman, R., & Squire, R.. Law and the rise of the firm. *Harvard Law Review*, **119**, 1335-1403 (2006).
3. Epstein, S. R. Craft guilds, apprenticeship, and technological change in preindustrial Europe. *The Journal of Economic History*, **58**(3), 684-713 (1998).
4. Padgett, J. F. & McLean, P. D. Organizational invention and elite transformation: The birth of partnership systems in renaissance florence. *American journal of Sociology* **111**(5), 1463-1568 (2006).
5. Chandler, A. D. *Scale and Scope: The Dynamics of Industrial Competition*. (Harvard Business School Press, 1990).

6. Simon, H. A. Organizations and markets. *Journal of Economic Perspectives*, **5**(2), 25-44 (1991).

7. Friedman, M. The social responsibility of business is to increase its profits. *New York Times Magazine*, 13 September, 122-126, (1970).

8. Freeman, R. E. *Strategic Management: A Stakeholder Approach*. (Cambridge university press, 2010).

9. Cavalli-Sforza, L. & Feldman, M. W. *Cultural Transmission and Evolution: A Quantitative Approach*, (Princeton University Press, 1981).

10. Boyd, R. & Richerson, P.J. *Culture and the Evolutionary Process*. (University of Chicago Press, 1985).

11. Boyd, R. & Richerson, P.J. *The Origin and Evolution of Cultures*. (Oxford University Press, 2005).

12. Mesoudi, A., *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. (University of Chicago Press, 2011).

13. Henrich, J. *The Secret of Our Success: How Culture is Driving Human Evolution, Domesticating our Species, and Making us Smarter*. (Princeton University Press, 2016).

14. Laland, K. *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. (Princeton University Press, 2017)

15. Smolla, M., & Akçay, E. Cultural selection shapes network structure. *Science Advances*, **5**(8), (2019).

16. Migliano, A.B. et al Hunter-gatherer multilevel sociality accelerates cumulative cultural evolution. *Science Advances*, **6**(9), (2020).

17. Cordes, C., Richerson, P. J., McElreath, R., & Strimling, P. A naturalistic approach to the theory of the firm: The role of cooperation and cultural evolution. *Journal of Economic Behavior & Organization*, **68**(1), 125-139, (2008).

18. Richerson, P. J., Collins, D., & Genet, R. M. Why managers need an evolutionary theory of organizations, *Strategic Organization*, **4**(2), 201-211, (2006)

19. Coase, R. The nature of the firm. *Economica* **4** (16), 386-405, (1937).

20. Hart, O., & Moore, J. Property rights and the nature of the firm. *Journal of Political Economy* **98** (6): 1119-1158, (1990).

21. Williamson, O. E. *Markets and Hierarchies*. (New York: Free Press, 1975).

22. Williamson, O. E. Comparative economic organization: The analysis of discrete structural alternatives. *Administrative Science Quarterly* **36** (2): 269-296, (1991).

23. Tadelis, S. Complexity, flexibility, and the make-or-buy decision. *American Economic Review*, **92** (2), 433-437, (2002).

24. Holmstrom, B. & Milgrom, P. The firm as an incentive system. *The American Economic Review* **84** (4), 972-991, (1994).

25. Lafontaine, F. & Slade, M. Vertical integration and firm boundaries: The evidence. *Journal of Economic Literature* **45** (3), 629-685, (2007).

26. Atalay, E., Hortaçsu, A. & Syverson, C. Vertical integration and input flows. *The American Economic Review* **104** (4), 1120-1148, (2014).

27. Argote, L. & Miron-Spektor, E. Organizational learning: From experience to knowledge. *Organization Science* **22** (5), 1123-1137, (2011).

28. Gibbons, R. & Henderson, R. Relational contracts and organizational capabilities. *Organization Science* **23** (5), 1350-1364, (2012).

29. De la Croix, D., Doepke, M. & Mokyr, J. Clans, guilds, and markets: Apprenticeship institutions and growth in the pre-industrial economy. *The Quarterly Journal of Economics*, **133** (1), 1-70, (2018).

30. Morrison AD & Wilhelm, WJ. The demise of investment banking partnerships: Theory and evidence. *The Journal of Finance*. **63** (1), 311-50, (2018).

31. Barnard, C. *The Functions of the Executive*. (Harvard University Press, 1938).

32. Schein, E. H. *Organizational Culture and Leadership* (4th edition). (John Wiley & Sons, 2010).

33. Podsakoff, N.P., Whiting, S.W., Podsakoff, P.M. & Blume, B.D., Individual-and organizational-level consequences of organizational citizenship behaviors: A meta-analysis. *Journal of applied Psychology*, **94** (1), 22, 2009

34. Kogut, B. & Zander, U. What firms do? Coordination, identity, and learning. *Organization Science*, **7** (5), 502-518, (1996).

35. Baker, G., Gibbons, R., & Murphy, K. J. Relational contracts and the theory of the firm. *The Quarterly Journal of Economics*, **117** (1), 39-84, (2002).

36. Fogarty, L., Strimling, P., & Laland, K.N. The evolution of teaching. *Evolution* **65** (10), 2760-2770, 2011.

37. Dean, L.G., Kendal, R.L., Schapiro, S.J., Thierry, B. & Laland, K. N. Identification of the social and cognitive processes underlying human cumulative culture. *Science* **335** (6072), 1114-1118, (2012).

38. Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. Cause and effect in biology revisited: is Mayr’s proximate-ultimate dichotomy still useful? *Science*, **334** (6062), 1512-1516, (2011).

39. Bateson, P. & Laland, K.N. Tinbergen's four questions: an appreciation and an update. *Trends in Ecology & Evolution*, **28** (12), 712-718, (2013).

40. Boyd, R. *A Different Kind of Animal: How Culture Transformed Our Species*. (Princeton University Press, 2018).

41. Mayr, E. Cause and effect in biology. *Science*, **134** (3489), 1501-1506, (1961).

42. Elster, J., 1983. *Explaining Technical Change: A Case Study in the Philosophy of Science*. (Cambridge University Press, 1983).

43. Nelson, R. R. & Winter, S. G. *An Evolutionary Theory of Economic Change*. (Belknap Harvard, 1982).

44. Dari-Mattiacci, G., Gelderblom, O., Jonker, J. & Perotti, E. The emergence of the corporate form *Journal of Law, Economics and Organization* **33** (2), 193-236, (2017).

45. Giuliano P, & Nunn, N. Understanding cultural persistence and change. Review of Economic Studies, (Forthcoming).

46. Gray, R.D. & Watts, J. Cultural macroevolution matters. *Proceedings of the National Academy of Sciences*, **114** (30), 7846-7852, (2017).

47. Coto-Sarmiento M., Rubio-Campillo, X. & Remesal, J. Identifying social learning between Roman amphorae workshops through morphometric similarity, *Journal of Archeological Science*, **96**, 117-123, (2018).

48. Smith, A. *The Wealth of Nations*. (W. Strahan and T. Cadell, London, 1776)

49. Becker, G. S. & Murphy, K. M. The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics*, **107** (4), 1137-1160, (1992)

50. Ricardo, D. *On the Principles of Political Economy and Taxation*. (London: John Murray, 1817).

51. Sterelny, K. *The Evolved Apprentice: How Evolution Made Humans Unique*. (MIT Press, 2012).

52. Rogers, A. R. Does biology constrain culture? *American Anthropologist* **90** (4), 819-831, (1988).

53. Boyd, R. & Richerson, P.J. Why does culture increase human adaptability? *Ethology and Sociobiology* **16** (2), 125-143, (1995).

54. Boyd, R., Richerson, P. J. & Henrich, J. The cultural evolution of technology: Facts and theories. In *Cultural Evolution: Society, Technology, Language, and Religion*, ed. Peter J. Richerson and Morten H. Christiansen, (2013).

55. Apel, J. Knowledge, know-how and raw material: The production of Late Neolithic flint daggers in Scandinavia. *Journal of Archaeological Method and Theory*, **15** (1), 91-111, (2008).

56. Smith, D., Stebbins, R., & Grotz, J. *The Palgrave Handbook of Volunteering, Civic Participation, and Nonprofit Associations*, (Pallgrave McMillan, 2016).

57. Anderson, R., Voluntary associations in history. *American Anthropologist*, 73, 209-222, (1971)

58. Lowie, R. *Social Organization*. (New York Rinehart and Co., 1948)

59. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314** (5805), 1560-1563, (2006).

60. Rand, D. G. & Nowak, M. A. Human cooperation. *Trends in Cognitive Sciences* **17** (8), 413-425, (2013).

61. Bowles, S. Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science* **324** (5932): 1293-1298, (2009).

62. Richerson, P. et al Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, 39, (2016).

63. Sober, E., & Wilson, D. S. Unto Others: The Evolution and Psychology of Unselfish Behavior. (Harvard University Press, 1999).

64. Nickerson, J., & Zenger, T. A knowledge-based theory of the firm -- The problem solving perspective. *Organization Science* **15** (6), 617-632, (2004).

65. Henrich, J. Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, 53(1), 3-35, (2004),

66. Levitt, S.D., List, J.A., & Syverson, C. Toward an understanding of learning by doing: Evidence from an automobile assembly plant, *Journal of Political Economy*, **121** (4), 643-681, (2013).

67. Buchanan, J. An economic theory of clubs, *Economica*, **32** (125): 1-14, (1965).

68. Anderson, G. M., Shughart, W. F., & Tollison, R. D. The economic theory of clubs. In *The encyclopedia of public choice*, 499-504, (Springer, Boston, MA, 2004)

69. Sandler, T. & Tschirhart, J. Club theory: Thirty years later. *Public choice*, 93(3-4), 335-355, (1997).

70. Arora, A., Cohen, W. & Cunningham, C. Inventive capabilities in the division of innovative labor, *NBER working paper No. 25051*, (2018).

71. Henderson, R. Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. *The RAND Journal of Economics*, 248-270, (1993)

72. Tripsas, M. & Gavetti, G. Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal*, **21** (10-11), 1147-1161, (2000).

73. Ward, Benjamin. The firm in Illyria: Market syndicalism. *American Economic Review*, **48** (4): 566–89, (1958).

74. Gustafsson, B. The rise and economic behavior of medieval craft guilds: An economic-theoretical interpretation. *Scandinavian Economic History Review*, **XXXV** (1): 1–40, (1987).

75. Spulber, D. F. *The Theory of the Firm: Microeconomics with Endogenous Entrepreneurs, Firms, Markets, and Organizations*. (Cambridge University Press, 2009)

76. Hoppitt, W. & Laland, K. N. Social Learning: An Introduction to Mechanisms, Methods, and Models. (Princeton University Press, 2013).

77. Cohen, W. M. & Levinthal, D. A. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 128-152, (1990)

78. Gross, D. The consequences of invention secrecy: Evidence from the USPTO patent secrecy program in world war II, *NBER Working Paper No. 25545*, (2020).

79. Kendal, J., Giraldeau, L. A., & Laland, K. The evolution of social learning rules: payoff-biased and frequency-dependent biased transmission. *Journal of Theoretical Biology*, **260** (2), 210-219, (2009).

80. Boyd, R., & Richerson, P. J. Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364** (1533), 3281-3288, (2009).

81. Turchin, P. *Ultra-Society: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth*. (Beresta Books, 2016).

82. Murdock, G. P. *Ethnographic Atlas*, (University of Pittsburgh Press, 1967).

83. Kirby, K. et al D-PLACE: A global database of cultural, linguistic and environmental diversity." *PloS one* **11** (7), (2016).

84. Oster, E. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, **37** (2), 187-204, (2016).

85. Comino, S., Galasso, A. & Clara Graziano, C. Market power and patent strategies: Evidence from renaissance Venice, *Journal of Industrial Economics*, forthcoming

9 Acknowledgements

Poblete gratefully acknowledges research support from ANID PIA/APOYO AFB180003 and Instituto Milenio MIPP IS130002. Both authors would like to acknowledge comments provided by Marcus Feldman, Peter Richerson, Christoph Loch, Stelios Kavadias, Brian Silverman, Joel Peck, Antoine Feylessoufi, Raicho Bojilov, Emilio Depetris-chauvin, Ricardo Alonso, Jose Tessada, Amir Goldberg, Sendil Ethiraj, Bart Vanneste, Vincent Mak and Tim Waring; by seminar participants at the Judge Business School, Stanford Business School, the London Business School; and by participants at the 2nd Annual Meeting of the Cultural Evolution Society, 7th Toulouse Economics and Biology Workshop, Annual SIOE Conference, Cambridge Clare Hall Graduate Students Conference, Third Culture Conference at the University of Birmingham, the INFORMS Annual Meeting, the Winter CJBS PhD Conference, the Sixth Development Workshop at the Catholic University of Chile, the inaugural Strategy Science Conference, Barcelona GSE Summer Forum (Organizational Economics) and the Trento Summer School on "New thinking on the firm: Perspectives from law, economics, organization, and history".

10 Author contributions

F.B. and J.P. conceived the study. J.P. did most of the mathematical models, reviewed the empirical analysis and reviewed and edited the text. F.B. reviewed and contributed with some of the mathematical models, executed the model simulations and associated figures, wrote most of the text, and executed the empirical analysis and its associated figures and tables.

11 Competing Interests

The authors declare no competing interests.

12 Figure legends

FIGURE 1. Equilibrium of the model for different values of λ.

We use C=0.6, c=0.45, $\widetilde{c}$=0.3, and p=0.1 to graph the equilibrium of the model for different values of λ. (A) Fitness inside and outside organization (we multiply fitness by 2.5 to obtain fitness equal or superior to 1). (B) Percentage of the population that has a technology tuned to the state. (C) Share of social and individual learners outside the organization. (D) Share of social and individual learners inside the organization.

FIGURE 2. Comparative statics.

For the baseline case, we reproduce the equilibrium fitness with PO depicted in panel A of figure 1. The red lines modify one parameter at the time. (A) We set c=0.4. (B) We set $\tilde{c}$=0.35. (C) We set p=0.12. (D) In the red line, we make the PO secretive, that is, social learners outside cannot imitate members of the PO (see the section "Robustness to different assumptions for social learning").

FIGURE 3. The impact of the presence of activities and POs on the size of local population.

The figures use the sample utilized in table 1, namely 173 societies. (A / B / C) Scatter plots in which the sizes of the bubbles represent the frequency of societies. (D) Here, we plot the average of marginal effects of the results displayed in the second column of table 1. We evaluate how much is the marginal effect of increasing the presence of activities increases from 0 to 1 on the probability of each one of the eight size categories (by definition, the changes add up to zero across all size categories). To explore how this impact varies with POs, we modify the value of "% within PO": i) in the blue line we set "% within PO" equal to zero (i.e., the societies go from zero to full adoption of technologies, but does so without using POs), ii) in the red line we set "% within PO" equal to 0.5 (i.e., the societies go from zero to full adoption of technologies, but does so without using POs). We display 95% confidence interval around the point estimate.

13 Tables

TABLE 1. Impact of presence of technologies and PO on the size of local population

	Dependent variable: Size of local population	
	β [t_{df}] (p-value) {95% CI}	
	1	2
% presence	1.137 [t_{114}=1.39] (0.165) {-0.466; 2.742}	0.016 [t_{113}=0.02] (0.984) {-1.641; 1.675}
% presence x % within PO		4.298 [t_{113}=3.36] (0.001) {1.729; 6.805}
Geographic controls?	Yes	Yes
Resource endowment controls?	Yes	Yes
Year of ethnography?	Yes	Yes
Agriculture intensity dummies?	Yes	Yes
Region dummies?	Yes	Yes
Type of settlement dummies?	Yes	Yes
Observations	173	173
Pseudo R Square	0.329	0.352

We execute ordered probit regressions. We assumed that the data, other than its categorical nature, met the assumptions of ordered probit regression (e.g., normality of errors). The dependent variable is the size of local population. Robust standard errors are used in all regressions. The estimated coefficient is presented without brackets. The t-test (with its corresponding degrees of freedom), the exact p-value, and the 95% confidence interval are reported in brackets.