Supplementary information for online publication

Index of SI

1 Mathematical proofs

Proof that there exists a unique ESS (r_I , r_S) and in that equilibrium, $f_S = (1 - C)$.

The proof has two parts: we first show that there exists a unique Nash equilibrium; moreover, in this equilibrium, $f_I = (1 - C)$. In the second part, we show that this equilibrium constitutes an ESS. There cannot exist other ESSs because the ESS is a refinement of the Nash equilibrium.

Lemma 1: *In any Nash equilibrium, both strategies are played.*

Proof. Suppose not and remember that $f_I = (1 - C)$. If $r_I = 0$, then $f_S = -c < f_S = 1 - C$. So r_I must be positive in any equilibrium. If $r_I = 1$, then $f_I = (1 - p) - c$. In order for f_S to be greater than $f_I = (1 - C)$, it is necessary that $(C - c) > p$. Since this condition will always be met according to assumption 1, then r_I can never be one because deviating to r_S would be profitable. ■

Lemma 2: *There is a unique Nash equilibrium* (r_I^*, r_S^*).

Proof. By lemma 1, in equilibrium, we must have $f_I = f_S$, but f_I is constant in r_I, and f_S is increasing in r_I. Therefore, there exists a unique r_I compatible with $f_I = f_S$, and thus, the equilibrium is unique. ■

Lemma 3: *The strategies played in the Nash equilibrium* (r_I^*, r_S^*) *are ESSs.*

Proof. Suppose that a population of size $\delta < 1$ with a share of individual learners $r_I \neq r_I^*$ invades. 1) If r_I is greater than r_I^*, then the resulting share of individual learners would be $\widetilde{r_I} = (1/(1+\delta))(r_I^* + \delta r_I) > r_I^*$. This result implies that in the new equilibrium the fitness of social learners is $\widetilde{f_S} > 1 - C$, because the fitness of social learners is increasing in the share of individual learners. The average fitness of the invading population would be $1 - C + (1 - r_I)(f_S - (1 - C))$, while the average fitness of the existing population would be $1 - C + (1 - r_I^*)(f_S - (1 - C))$, which is strictly greater because $r_I > r_I^*$. 2) If r_I^* is greater than r_I, then the resulting share of individual learners would be $\widetilde{r_I} = (1/(1+\delta))(r_I^* + \delta r_I) < r_I^*$. This result implies that in the new equilibrium, the fitness of social learners is $\widetilde{f_S} < (1 - C)$ because the fitness of social learners is increasing in the share of individual learners. The average fitness of the invading population would be $1 - C - (1 - r_I)((1 - C) - \widetilde{f_S})$, while the average fitness of the existing population would be $1 - C - (1 - r_I^*)((1 - C) - \widetilde{f_S})$, which is strictly less because r_I<r_I^*}. ■

Because every set of ESSs is Nash, we know there exists a unique set of ESSs for this game. QED

Proof of Proposition 1. The structure of the proof is identical to the previous one, including a PO of size λ, for a sufficiently small λ.

Lemma 4: *In any Nash equilibrium, both strategies are played outside the PO if* λ is sufficiently small..

Proof. Suppose not and remember that $f_I = (1 - C)$. If $r_I = 0$, then $f_S = \lambda\tilde{q} - c < f_I = 1 - C$ if λ is sufficiently small. So, r_I must be positive in any equilibrium. If $r_I = 1$, then $f_S = \lambda(q - c) + (1 - \lambda)(1 - c)(1 - p)$. In order for f_S to be greater than $f_I = (1 - C)$ for every $\lambda > 0$, it is necessary that $(C - c) > p - pc$. Since this condition will always be met according to assumption 1, r_I can never be one because deviating to r_S would be profitable. ■

Lemma 5: *There is a unique Nash equilibrium* (r_I^*, r_S^*) *outside the PO.*

Proof. By lemma 1, in equilibrium, we must have $f_I = f_S$, but f_I is constant in r_I and f_S is increasing in r_I; therefore, there exists a unique r_I compatible with $f_I = f_S$, and thus, the equilibrium is unique. ■

Lemma 6: Inside the PO, all agents play "social learning".

Proof. The fitness of individual learning is $\widetilde{f_I} = 1 - C,$ and the fitness of social learning inside the PO is given by $\widetilde{f_S} = f_S + \lambda(c - \tilde{c}) = f_I + \lambda(c - \tilde{c})$ (see equation 2 in the main body). Because $\widetilde{f_I} < \widetilde{f_S}$, only social learning is played inside the firm. ■

Lemma 7: *The strategies played in the Nash equilibrium* $(r_I^*, r_S^*, \widetilde{r_S} = 1)$ *are ESSs.*

Proof. Suppose a population of size $\delta < 1$ invades. 1) If the invasion is outside the PO, the proof is analogous to that of lemma 3 of the previous proof and therefore omitted. If the invasion is inside the PO and $r_I \neq 0$, observe that average fitness of invading population is $r_I(1-C)+(1-r_I)f_S < f_S$, and thus, it has lower fitness than the population inside the PO. ■

Because every set of ESSs are Nash, we know there exists a unique set of ESSs for this game. QED

Proof that the size of the PO that maximizes society's fitness is higher than the size that maximizes fitness of the average member of the PO.

Proof. Let $\widetilde{f}\,(\lambda)$ be the fitness of agents inside the PO. At the level that maximizes fitness, it must be that $\widetilde{f}\,'(\lambda) = 0$. The overall fitness of the population is given by $(1-C)(1-\lambda)+\lambda\,\widetilde{f}\,(\lambda)$; therefore, the derivative with respect to the size λ is $-(1-C)+\widetilde{f}\,'(\lambda)$ which at the point that maximizes member's fitness is negative. By lemma 6 we know that $\widetilde{f}\,(\lambda) = (1-C)+\lambda(c-\widetilde{c}\,)$ and $\widetilde{f}\,'(\lambda) > 0$. Therefore, the society will benefit from increasing the size of the PO further. QED

Proof of that proposition 1 is robust to PO biased social learning, provided that the bias is small enough.

It is reasonable to assume that social learner inside the PO, given that they enjoy a lower social learning costs, will not learn randomly and will privilege, instead, learning socially from fellow PO members. To model this, first we modify the recursion for the tuned population inside the PO (outside it remains the same):

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

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

The parameter α denotes the extent of the PO bias inside the PO and satisfies $0 \leq \alpha \leq 1-\lambda$.

From these equation we find the steady state values for

$$q^e = \frac{r_I + (1-p)\,\lambda\,(r_S - \widetilde{r_S}) - r_I\widetilde{r_S}(1-p)\alpha}{1-(1-p)[(\lambda+\alpha)\widetilde{r_S} + (1-\lambda)\,r_S - \widetilde{r_S}r_S(1-p)^2\alpha]}$$

$$\tilde{q}^e = \frac{\widetilde{r_I} + (1-p)(1-\lambda)(\widetilde{r_S} - r_S) - r_I\widetilde{r_S}(1-p)\alpha}{1-(1-p)[(\lambda+\alpha)\widetilde{r_S} + (1-\lambda)\,r_S - \widetilde{r_S}r_S(1-p)^2\alpha]}$$

If $\alpha = 0$ we return to the equations (3) and (4) of the main body of the manuscript. While the fitness expression of social learners outside the firm remains the same, the expression for those inside the PO changes:

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

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

And thus,

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

It is trivial to show that proposition 1 will still hold provided that α is small enough. For any λ for which proposition 1 holds without PO bias, it is possible to find a α small enough so that $\widetilde{f_S} > f_S$ under PO bias. QED

Simulation. It interesting to explore how $\widetilde{f_S}$ behaves for different values of α. If α increases, two effects takes place. On the one hand, the PO will benefit from the lower social learning cost (i.e., positive term $c - \tilde{c}$). However, on the other hand it will suffer from being increasingly less tuned to the environment (i.e., negative term $\tilde{q}^e - q^e$). This can be verified by inspecting the following expression:

$$\tilde{q}^e - q^e = \frac{p\,(\widetilde{r_I} - r_I)}{1-(1-p)[\lambda\widetilde{r_S} + (1-\lambda)\,r_S + \alpha\widetilde{r_S}(1-r_S(1-p)^2)]}$$

Assuming that α is small enough so that $\widetilde{f_S} > f_S$, and therefore $\widetilde{r_I} = 0$, it is easy to verify that the negative value of $\tilde{q}^e - q^e$ will increase if α grows.

To explore in detail how $\widetilde{f_S}$ behaves, we execute a simulation. We assume that α is small enough so that $\widetilde{f_S} > f_S$, and then, by replacing q^e and $\tilde{q}^e$ in the fitness expression outside the PO and equalizing it to $1 - C$ (i.e., an internal equilibrium occurs outside), we find the following expression for r_I,

$$r_I = \frac{[1-\alpha(1-\mathrm{p})]}{[(1-\lambda)-\alpha(1-\mathrm{p})]}\cdot\frac{p}{(1-p)}\cdot\frac{[1-(C-c)]}{(C-c)}$$

If $\alpha = 0$ we return to the value of r_I displayed in the main body of the paper.

With these expressions it is straightforward to produce supplementary figure 2 (see the end of the SI for the figures and tables of this appendix). In this figure we vary α and plot $\widetilde{f_S}$, the fitness of social learners inside the PO. Consider the case of $\lambda = 0.1$ (with this size of the PO, $0 \le \alpha \le 0.9$). We find that, initially, increasing α increases the fitness of social learners which, given that $\widetilde{r_S} = 1$, correspond to the fitness of the PO. However, after a peak is reached at approximately $\alpha \approx 0.33$, further increases in α generate a reduction in fitness. Moreover, we find that after $\alpha \approx 0.56$, the fitness of a PO populated only by social learners drops below 1, and thus i) the PO is no longer beneficial to society (its fitness will be equal to

$1 - C$) and ii) an internal equilibrium inside the PO will occur (i.e., there will a positive share of individual learners inside the PO). The case for λ equal to 0.2 and 0.3 yield qualitatively similar results. However, for the cases λ equal to 0.4, 0.5 and 0.6, we find that fitness always decreases as α grows. This is intuitive: when a PO populated by social learners is smaller, the penalty of being less tuned to the environment decreases (this can be verified by perusing the expression for $\tilde{q}^e - q^e$ above), and thus, the gains in lower social learning by expanding α can be larger than the drop in environmental tuning.

We also produced supplementary figure 3, where we explore, for a given small size of $\lambda = 0.1$, two cases: when the advantage in the cost of social learning is small (i.e., $c - \tilde{c} = 0.05$) versus when it is large (i.e., $c - \tilde{c} = 0.3$). In line with intuition, we find that when the advantage is small, the small size PO doesn't display the inverted U between α and fitness; instead the advantage is always decreasing and gets eliminated fairly quickly, around $\alpha \approx 0.15$.

Considering these simulations results, one could infer that evolution might favor a PO bias in social learning within POs, but only small ones, and provided that the advantage in social learning is large enough. Once PO reach a certain size, or if the advantage is small, random social learning would be a superior strategy vis-à-vis PO bias.

The fact that the PO bias in social learning has to be limited is consistent with empirical evidence and conventional wisdom on POs. In sociology, it is well known that organizations cannot be exclusively inwardly focused and need to be open to the environment (see Scott, 2003). The management literature shows that absorbing knowledge from the external environment is important for the success of firms (Argote and Miron-Spektor 2011; Cohen and Levinthal, 1990). The guilds literature also showcases the importance of the journeymanship (Epstein, 1998; De la Croix et al, 2018), which dictated that "after their training in a local guild, new artisans would travel to another city to acquire additional skills before they would qualify as masters" (De la Croix et al, 2018; p. 14).

Proof that secrecy reduces the fitness of PO's

Secrecy means that social learners outside the PO cannot imitate members of the PO. Consequently, outside the PO, we are back to the original case without a PO, in which the fitness of social learners outside the PO is

$$f_S = (1 - p)q^e - c \;, \tag{3}$$

independent of the size of the PO. The fitness of social learners inside the PO is as before, given by

$$\widetilde{f_S} = (1-p)[(1-\lambda)\cdot q^e + \lambda \cdot \tilde{q}^e] - [(1-\lambda)c + \lambda\tilde{c}]. \quad (4)$$

In this case, it is not straightforward whether f_S or $\widetilde{f_S}$ is higher.

To prove that proposition 1 is robust to secrecy, consider two societies with a productive organization of size λ, and assume that society a has secrecy and society b does not. Inside the PO (both with and without secrecy), all agents are social learners, and therefore, in either case, their fitness is given by

$$\widetilde{f}_S^{\,i} = [1-\lambda][(1-p)q^i - c] + \lambda[(1-p)\widetilde{q}^{\,i} - \widetilde{c}] \quad (8)$$

The recursive equations that determine the stock of knowledge inside the firm in the steady state is the same regardless of whether there is secrecy and is determined by

$$\widetilde{q}^{\,i}(t) = [\lambda\widetilde{q}^{\,i}(t-1) + (1-\lambda)q^i(t-1)](1-p) \quad (9)$$

From equations (8) and (9), it is clear that fitness in the firm is strictly increasing in the stock of knowledge outside the firm. Therefore, to show that secrecy is detrimental to the firm's fitness, we only need to show that secrecy is detrimental to the stock of knowledge outside the firm (i.e., that $q^a > q^b$).

To show this, it is useful to define the function $\widetilde{q}^{\,i}(q^i)$, which is the stock of knowledge inside the firm. This function is increasing in the stock of knowledge outside the firm, where $\widetilde{q}^{\,i}\left(q^i\right) < q^i$.

We know that

$$f_S^a = (1-p)q^a - c = 1 - C$$

$$f_S^b = (1-\lambda)(1-p)q^b + \lambda(1-p)\widetilde{q}^{\,b}(q^b) - c = 1 - C$$

In contrast, suppose that $q^a > q^b$; then,

$$1 - C = (1-p)q^a - c > (1-\lambda)(1-p)q^b + \lambda(1-p)\widetilde{q}^{\,b}(q^b) - c = 1 - C$$

is a contradiction. Therefore, it must be that $q(b) > q(a)$. QED

Proof of Proposition 3. Observe that the fitness of a social learner that specializes in technology j can be expressed as

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

Remember that in equilibrium, f_S^j and x^j must be constant across J outside the productive organization. It is straightforward to check that there are two Nash equilibria inside the PO. 1) Either $\tilde{x}^j$ is the same for every j in J, in which case $\tilde{f}_S^j$ is constant in j, or 2) there is full specialization, and $x^j = 1$ for $j = j^*$ and 0 otherwise.

Observe now that the first equilibrium is not ESS because if a small population of size δ that specializes in one activity j invades the PO, then the fitness of the invading population becomes

$$\tilde{f}_S^{\tilde{J}} = f_S^j + (c - \tilde{c}) \cdot \left(\frac{\lambda\, \tilde{x}^j + \delta}{\lambda\, \tilde{x}^j + (1-\lambda)x^j + \delta}\right)$$

The average fitness of the average population is

$$E_{j\in J}\left(\tilde{f}_S^j\right) = f_S^j + (c - \tilde{c}) \cdot \left[\frac{J-1}{J}\right]\left[\left(\frac{\lambda\, \tilde{x}^j}{\lambda\, \tilde{x}^j + (1-\lambda)x^j + \delta}\right)\right] + \left[\left(\frac{1}{J}\right)\right]\tilde{f}_S^{\tilde{J}} \; < \tilde{f}_S^{\tilde{J}}$$

To see that a specialized Nash is ESS, notice that in a specialized equilibrium, the fitness inside the PO is given by

$$\tilde{f}_S^{j^*} = f_S^{j^*} + (c - \tilde{c}) \cdot \left(\frac{\lambda}{\lambda + (1-\lambda)x^j}\right)$$

Observe that an invasion of size δ by any other technology will obtain a fitness of

$$\tilde{f}_S^{\tilde{J}} = f_S^j + (c - \tilde{c}) \cdot \left(\frac{\delta}{\lambda + (1-\lambda)x^j + \delta}\right)$$

which converges to $f_S^j < \tilde{f}_S^{\tilde{J}}$ as δ approaches 0, and therefore, the invading population obtains a lower fitness level. QED

Proposition 3 can be extended to include multiple PO's in the society. Suppose there are N POs, each of size λ. If the product N times λ if sufficiently small, then the equilibrium outside the PO's have both individual and social learners of every technology by the same argument as proposition 1. Then the argument presented in proposition 3 holds for every PO, and therefore all of them must be specialized, obtaining a fitness larger than the fitness of agents outside PO's.

2 A model of selective learning

In every period, a continuum of long-lived agents adopts 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 fitness. 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. Agents adopt a technology by using one of two behavioral strategies: learning by themselves at a cost C, labelled "individual learning", or copy from others of the previous period at a cost c ($c < C$), labelled "social learning". They choose between these two strategies selectively after observing a signal x that is drawn from distribution that depends on whether the state of the world has changed. (This is a standard way of modeling selective learning in the literature, see Boyd and Richerson (2005; chapter 1) and Boyd (2018; chapter 1).) If the state of the world has not changed, then the signal is distributed exponential with parameter $\alpha_0 = 1$, if the state of the world changes then the distribution is exponential with parameter $\alpha_1 < 1$. The cumulative and density function are described by F_0 and F_1, and, f_0 and f_1, respectively.

As is standard in Bayesian models the optimal strategy is a cutoff strategy in which the agent will choose to investigate the state by herself if the signal is above a threshold value d. The value of d will be considered a cultural trait that is inherited and that evolution will affect so that fitness is maximized. As in Boyd and Richerson (2005; chapter 1), this is translated into an ESS equilibrium in which the population has a cutoff d^* and a small mutant population with parameter $d^* + \varepsilon$ can't invade the population when ε is sufficiently small. In simple terms, the equilibrium d^* is the point where moving the threshold d by a tiny bit doesn't pay.

Let q be the fraction of the tuned population in the model, then the expected fitness of an individual with parameter d if the world remains unchanged, will be

$$h_u(d) = F_0(d)(q - c) + (1 - F_0(d))(1 - C)$$

If the world changes, the fitness will be

$$h_c(d) = F_1(d)(-c) + (1 - F_1(d))(1 - C)$$

The expected fitness will therefore be

$$h(d) = ph_c(d) + (1 - p)h_u(d)$$

The expected fitness of a mutant will be

$$h_u(d + \varepsilon) = h_u(d) + [F_0(d + \varepsilon) + F_0(d)](q - c - 1 + C)$$

$$h_c(d + \varepsilon) = h_c(d) + [F_1(d + \varepsilon) + F_1(d)](-c - 1 + C)$$

$$h(d + \varepsilon) = ph_c(d + \varepsilon) + (1 - p)h_u(d + \varepsilon)$$

The equilibrium condition can therefore be stated as follows:

$$\frac{\partial}{\partial \varepsilon} h(d+\varepsilon)|_{\{\varepsilon=0\}} = 0$$

$$(1-p)[f_0(d)](q-c-1+C) + p[f_1(d)](-c-1+C) = 0 \qquad (1)$$

Observe that in equilibrium it must be the case that $q-c-1+C>0$, because the term $-c-1+C$ is always negative. Thus, the equilibrium of d is reached when, after increasing d by a tiny bit, the expected gains of copying when the world doesn't change is equal to the expected loss of copying when it does.

The fraction q of expected tuned individuals is governed by the following recursion:

$$q(t+1) = (1-p)(F_0(d)\cdot q(t) + (1-F_0(d))) + p(1-F_1(d))$$

Therefore, the expected ratio of tuned population in equilibrium is,

$$q^e = \frac{(1-p)(1-F_0(d)) + p(1-F_1(d))}{1-(1-p)F_0(d)} \qquad (2)$$

Using equations (2) and (1) one can find the equilibrium values of d^* and q^e. Most importantly, in this model d^* – which determines the share of social learners through $F(d)$ – is decreasing in the cost of social learning c. This is intuitive: the smaller c is, the more convenient it is to learn socially (i.e., a higher d). Thus, we introduce the following proposition:

Proposition S1: *The equilibrium level of d (that determines the prevalence of social learning) is decreasing in c.*

Proof of proposition S1. This can be obtained directly from replacing (2) in (1), obtaining d^* and taking the derivative. To avoid excessive algebra this can also be done using the implicit function theorem. Using this theorem we can write equation (1) in terms of $d^* = \varphi(c,q)$, where φ is a differentiable function and then obtain,

$$\frac{\partial \varphi(c,q)}{\partial q} = \frac{-(1-p)f_0(d)}{(1-p)[(1-p)[f_0{}'(d)](q-c-1+C) + p[f_1{}'(d)](-c-1+C)]}$$

Because the distributions are exponential $f_0{}'(d) = -f_0(d)$ and $f_1{}'(d) = -\alpha_1 f_1(d)$ and therefore

$$\frac{\partial \varphi(c,q)}{\partial q} = \frac{-(1-p)f_0(d)}{(1-p)[(1-p)[f_0(d)](q-c-1+C) - p[\alpha_1 f_1(d)](-c-1+C)]}$$

Because $\alpha_1 < 1$, the denominator can be rewritten as

$$-\{(1-p)[f_0(d)](q-c-1+C) + p[f_1(d)](-c-1+C)\} - p[f_1(d)](1-\alpha_1)(-c-1+C)]$$

In the neighborhood around the equilibrium (in which equation (1) holds) the term in brackets {.} is zero, and thus we have that $\frac{\partial \varphi(c,q)}{\partial q} < 0$.

Using the implicit function theorem again, we obtain,

$$\frac{\partial \varphi(c,q)}{\partial c} = \frac{-(1-p)f_0(d)}{(1-p)[(1-p)[f_0'(d)](q-c-1+C) + p[f_1'(d)](-c-1+C)]}$$

Using the analogous argument as before, we can show that $\frac{\partial \varphi(c,q)}{\partial q} < 0$.

Equation (2) can be written as $q = \eta(d)$ where η is a continuous function. It is easy to verify that $(\partial \eta(d))/(\partial d) < 0$. Thus, the equilibrium level of d can be written as $d^* = \varphi(c, \eta(d))$. Differentiating d^* with respect to c we obtain the following,

$$\frac{\partial d^*}{\partial c} = \frac{\partial \varphi(c,\eta(d))}{\partial c} + \frac{\partial \varphi(c,\eta(d))}{\partial \eta(d)} \cdot \frac{\partial \eta(d)}{\partial d}$$

And because both terms are negative, we know that $\frac{\partial d^*}{\partial c} < 0$. ■

Notice that in this model the fitness at equilibrium increases if c becomes smaller, that is, it can be verified that the derivative $dh(d^*, q^e)/dc$ is negative. Thus, unlike the basic model of non-selective learning, where social learners generate a negative externality – a lower social cost yields the expansion of social learner but do not increase fitness –, here decreases of social learning costs are useful for society.

Finally, as highlighted in the literature (Boyd et al, 2013; Laland 2017) and consistent with the previous paragraph, observe that in this model Roger's paradox does not hold. This is because the fitness of the individuals using selective learning is larger than that of individuals doing pure individual or social learning strategies. It is easy to show that $h(d)$ is always higher than $1 - C$: the inequality simplifies to $p(c(F_0 - F_1) + (1 - C)\ (F_0 - F_1) + F_0(q(1-p) - 1 + C - c) > 0$, which is always true given that the implicit assumption $C - c > p$ (which allows a pure social learning strategy to evolve; see the baseline model).

Selective learning with a PO

We now introduce a PO in the model of selective learning. To simplify the discussion we make the following assumption: i) people outside the PO only copy from people outside the PO and, ii) people inside the PO copy randomly from the entire population. Assumption i) correspond to the case of secrecy in our baseline model and is necessary to simplify the mathematical analysis, as we only need to include equations from individuals inside the PO. Recall that secrecy did not affected qualitatively the results in the baseline analysis, and thus, using a selective learning model with secrecy is valid for comparison purposes.

Assume a fraction λ of the population is in the PO. Let q^* and d^* be the equilibrium values outside the PO (which is not directly affected by the PO). As before, the cost of social learning within the PO is $\tilde{c} < c$. Let $\tilde{q}$ and $\tilde{d}$ be the parameters inside the PO.

The expected fitness in the PO will be,

$$\widetilde{h_u}(\tilde{d}) = F_0(\tilde{d})(\lambda\tilde{q} + (1-\lambda)q^* - (\lambda\tilde{c} + (1-\lambda)c)) + (1 - F_0(\tilde{d}))(1 - C)$$

If the world changes, it will be

$$\widetilde{h_c}(\tilde{d}) = F_1(\tilde{d})(-(\lambda\tilde{c} + (1-\lambda)c)) + (1 - F_1(\tilde{d}))(1 - C)$$

The expected fitness will therefore be

$$\tilde{h}(\tilde{d}) = p\widetilde{h_c}(\tilde{d}) + (1-p)\widetilde{h_u}(\tilde{d})$$

The expected fitness of a mutant will be

$$\widetilde{h_u}(\tilde{d} + \varepsilon) = \widetilde{h_u}(\tilde{d}) + [F_0(\tilde{d} + \varepsilon) + F_0(\tilde{d})]((\lambda\tilde{q} + (1-\lambda)q^* - (\lambda\tilde{c} + (1-\lambda)c) - 1 + C)$$

$$\widetilde{h_c}(\tilde{d} + \varepsilon) = \widetilde{h_c}(\tilde{d}) + [F_1(\tilde{d} + \varepsilon) + F_1(\tilde{d})](-(\lambda\tilde{c} + (1-\lambda)c) - 1 + C)$$

$$g(\tilde{d} + \varepsilon) = p\widetilde{h_c}(\tilde{d} + \varepsilon) + (1-p)\widetilde{h_u}(\tilde{d} + \varepsilon)$$

As before, the equilibrium condition is the following:

$$\frac{\partial}{\partial\varepsilon} h(\tilde{d} + \varepsilon)|_{\{\varepsilon=0\}} = 0$$

$$(1-p)[f_0(\tilde{d})]((\lambda\tilde{q} + (1-\lambda)q^* - (s\tilde{c} + (1-\lambda)c) - 1 + C) + p[f_1(\tilde{d})](-(\lambda\tilde{c} + (1-\lambda)c) - 1 + C) = 0 \quad (3)$$

The fraction $\tilde{q}$ of expected tuned individuals inside the PO is governed by the following recursion:

$$\tilde{q}(t+1) = (1-p)(F_0(\tilde{d}) \cdot (\lambda\tilde{q} + (1-\lambda)q^*) + (1 - F_0(\tilde{d}))) + p(1 - F_1(\tilde{d}))$$

Therefore, the expected ratio of tuned population inside the PO in equilibrium is,

$$\widetilde{q^e} = \frac{p(1 - F_1(\tilde{d})) + (1-p)(1 - F_0(\tilde{d}) + F_0(\tilde{d})(1-\lambda)q^*}{1 - (1-p)F_0(d)} \quad (4)$$

Using equations (2) and (1) one can find the equilibrium values of $\widetilde{d^*}$ and $\widetilde{q^e}$. Several important results emerge from this analysis:

First, notice that using the same argument as before, proposition S1 also holds for the PO equilibrium as well: the amount of social learning is decreasing in $\widetilde{c}$.

Second, as a corollary of the first result, observe that, given that inside the PO have a lower average cost of social learning $\widetilde{c} < c$, individuals will choose $\widetilde{d^*} > d^*$, for any s. This means that, just as in our model without selective learning, social learning is more prevalent inside the PO, while individual learning is more prevalent outside. Here the difference is less dramatic because there always exists some individual learning inside the PO.

Third, also notice that while the fitness outside the PO doesn't change, it is possible to verify that the derivative $d\widetilde{h}\,(\widetilde{d^*},\widetilde{q^e})/d\widetilde{c}$ is negative, and thus, that the fitness of the PO at equilibrium increases if $\widetilde{c}$ becomes smaller. Thus, similarly to our baseline model, the introduction of a PO increases the fitness of the population. However, given that Rogers' paradox is not present in this model, the PO enhances the capacity of social learning to generate cumulative culture. In contrast, in the non-selective model, the PO is a sufficient condition.

Fourth, the dynamics when the PO grows are the same as in our model without selective learning. When the PO grows, two effects are in place. First, the likelihood of copying from other people inside the PO increases, and thus individuals experience a reduction in the average cost of social learning, given by $\lambda\widetilde{c} + (1-\lambda)c$. Second, agents inside the PO will evolve to increase their critical level d to take advantage of this lower cost (see proposition S1). This creates a negative externalities on other PO members because now, every time they choose to learn socially, they are more likely to earn from social learners that are less likely to be tuned to the state.

However, it is important to notice that in the selective learning model, given that social learners do not generate a negative externality (see discussion above), simulations we performed show that it is not certain that the second effect will eventually outweigh the first effect. Thus, depending on the parameters, the PO might evolve to include the whole population and the PO doesn't need to be small to benefit society. We interpret this result in two ways. First, it shows that selective learning enlarges the proposition 1 of the paper. While the non-selective model requires λ to be small, in a selective learning model this is not necessary. This may enlarge the type of organizations that the theory can encompass: while the small size condition applies well to "productive organizations" such as guilds and firms (which tend to be small in relation to the population), other organizations are very large such as many public services (e.g., the military, health, education, religion). Second, it is likely that the advantage of social learning inside the PO is reduced when its size increases, this assumption is well studied and documented in the management and economics literature: firms display diseconomies of scale and cooperation tends to falter with size (Brahm and Tarzijan, 2012; Coase, 1937; Shaver and Mezias, 2009; Zenger et al, 2011; Cordes et al; 2008). In the extension “Multiple POs, cooperation and group selection” below, we formalize this idea: cooperation inside the PO, and thus the extent of its social learning advantage, is reduced with the size of the PO.

Summary. A summary of the role of the PO in unbiased vs. selective social learning environments is in order. While in selective learning the role of the PO is not solving Rogers’ paradox, and thus the PO is not crucial for cumulative culture –the paradox is solved by selective learning itself–, all the other results remain largely unchanged: the PO enhance the fitness of the society (i.e., it further propels cumulative culture), the proportion of social learner 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 increasing exposure to an environmental change.

3 A model of payoff biased social learning

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 fitness. 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. 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 member of the population did in the previous period and simply copies its technology, incurring a cost $c < C$. We assume that social learners are "payoff biased". Every period a proportion ϕ of social learners is capable of observing the fitness of the rest of the population and copy the group that displays the highest fitness. In our case, these groups are two: fellow social learners or individual learners. For example, if a change in environment just occurred, individual learners would display higher fitness and thus social learner would copy them rather than other social members. Or if social learners are just a few, and there are many individual learners, most likely social learners would display higher fitness and they would be copied. With the remaining portion of $(1 - \phi)$ social learners simply copy unbiasedly, choosing a random individual in the society. Thus, if $\phi = 0$ we are back to the main model in the manuscript.

Let r_I be the share of individual learners and r_S the share of social learners in the population. Let $q(t)$ be the percentage of individuals in the population with a tuned technology in period t and $q_S(t)$ be the percentage of social learners with a tuned technology in period t. Because individual learners are always tuned, the fraction of tuned individuals in the population is,

$$q(t) = r_I + r_S \cdot q_S(t)$$

Let $f_S(t)$ be the average fitness of social learners in period t. The average fitness of social learners at t is given by $f_S(t) = q_S(t) - c$. The fitness of individual learners is always $f_I(t) = 1 - C$. Therefore, in each period social learners enjoy a larger fitness on average if,

$$q_S(t) > 1 - C + c$$

The fraction of tuned social learners in period $t + 1$ will depend on whether the state has changed or not. Let J be a variable that take value 1 if the state of the world has change and 0 otherwise (whose mean is p). The recursion equation that determines $q_S(t + 1)$ is the following,

$$q_S(t+1) = \left\{ \begin{array}{cc} 0 & if\ J = 1 \\ [\phi + (1-\phi)q(t)] & if\ J = 0\ and\ q_S(t) < 1 - C + c \\ [\phi \cdot q_S(t) + (1-\phi)q(t)] & if\ J = 0\ and\ q_S(t) \geq 1 - C + c \end{array} \right\}$$

The model cannot be approximated linearly as before, and therefore we analyze this model using simulations. We initialize the simulation seeding values of q_S and r_I and then we let it run for 150,000 periods (assuming a standard replicator dynamic equation and weak selection). We compute the average between periods 50,001 and 150,000 as the "long run" "equilibrium level" of the variables of interest. Supplementary figures 4 and 5 describe the details of the simulation and the results we obtain (see the "Figures and tables" section of this appendix below). The simulation provides three main findings. (We tried many different values of C, c and p, and the results remained qualitatively the same.)

First, we find that the equilibrium shares of social (individual) learners increases (decreases) with the extent of payoff bias. This result is intuitive as payoff bias makes social learners more effective in copying.

Second, we find that the fitness of social and individual learners equalize in the long run (see supplementary figures 4 and 5). This means that Roger's paradox holds for payoff bias. This is intuitive and aligned with the literature: payoff bias doesn't help individual learners to be more frequent or more effective in their learning (Boyd and Richerson, 2005; chapter 2). Thus, a role of the PO might well be to overcome this paradox even in the presence of payoff bias. (Notice, however, that payoff bias does break Roger's paradox for more than one technology[1].)

Third, we find that, consistently with the non-linearity of the recursion of q_S, the equilibrium level of the share of social learners also display non-linear behavior (see supplementary figure 5). Given our parameters we observe a change in level that occurs around a bias of 0.74. Fully explaining this non-linear behavior is beyond the scope of this article.

Notice that in supplementary figure 5, as in the remaining figures of this model below, we display only values up to a bias of 0.9. This is because for very high payoff bias a steady state equilibrium may fail to exist. To understand why consider a payoff bias of 1. In that case, as long as there is a positive measure of individual learners all social learners will copy from them and obtain a larger fitness. Replication dynamics means that social learners might invade the population with individual learners disappearing altogether. However, if the environment changes, a mutation to individual learners will now obtain a larger fitness. But after the change, social learners would then thrive, invading again. This cycle would continue and no meaningful integer number of individual learner might be computed in expectation, and thus, we might fail to have a steady state equilibrium. As before, a careful analysis of what happens for larger biases is beyond the scope of this paper.

Productive organizations

We now introduce PO into the model. As in the baseline model, a fixed fraction of agents λ is located inside the organization, and social learners inside the PO enjoys a lower cost $\tilde{c}$ if they learn from the inside. As before, we define the variables inside the PO as $\tilde{r}_I$, $\tilde{r}_S$, $\tilde{q}$ and $\tilde{q}$. To save on notation we define the tuned share for the entire population,

$$q_T(t) = (1 - \lambda)\tilde{q}(t) + \lambda\tilde{q}(t)$$

[1] When there are multiple traits that individuals use, and social learner may copy different traits from different individuals, payoff biased social learning will generate an increase in fitness due to the creation of novelty (see Boyd et al, 2013 and Boyd and Richerson, 1985, chapter 8).

As before, $f_S(t) = q_S(t) - c$. The average fitness of a social learner inside a PO depends on whom they copy from. Let $\beta(t)$ be the chance they copy from other PO members in period t then

$$\widetilde{f_S}(t) = \widetilde{q_S}(t) - c + \beta(t)(c - \tilde{c})$$

As before, social learners enjoy a larger fitness on average if $q_S(t) > 1 - C + c$. Inside the PO, the fitness of social learners is larger than that of social learners outside if

$$\widetilde{q_S}(t) + \beta(t)(c - \tilde{c}) > q_S(t)$$

The fraction of tuned social learners outside the PO in period $t + 1$ will depend on whether the state changes and on the level $q_S(t)$. The recursion equation that determines $q_S(t + 1)$ is determined as follows

$$q_S(t+1) = \left\{ \begin{array}{cc} 0 & if\ J = 1 \\ [\phi + (1-\phi)q_T(t)] & if\ \ J = 0\ and \max\{1 - C, \widetilde{f_S}(t), f_S(t)\} = 1 - C \\ [\phi \cdot q_S(t) + (1-\phi)q_T(t)] & if\ \ J = 0\ and \max\{1 - C, \widetilde{f_S}(t), f_S(t)\} = f_S(t) \\ [\phi \cdot \widetilde{q_S}(t) + (1-\phi)q_T(t)] & if\ \ J = 0\ and \max\{1 - C, \widetilde{f_S}(t), f_S(t)\} = \widetilde{f_S}(t) \end{array} \right\}$$

The variable $\widetilde{q_S}(t + 1)$, that is, the fraction of tuned social learners inside the PO in period $t + 1$, is determined by the same equation as $q_S(t + 1)$.

Finally the fraction $\beta(t)$ is determined as follows,

$$\beta(t+1) = \left\{ \begin{array}{cc} \phi \dfrac{\lambda \tilde{r}_I}{\lambda \tilde{r}_I + (1-\lambda) r_I} + (1-\phi)\lambda & if \max\{1 - C, \widetilde{f_S}(t), f_S(t)\} = 1 - C \\ (1-\phi)\lambda & if \max\{1 - C, \widetilde{f_S}(t), f_S(t)\} = f_S(t) \\ \phi + (1-\phi)\lambda & if \max\{1 - C, \widetilde{f_S}(t), f_S(t)\} = \widetilde{f_S}(t) \end{array} \right\}$$

In this model, if ϕ and λ are equal to zero, we converge to our baseline model. As before, we analyze the model using simulations. The simulation are analogous to the ones made in the case without PO. The supplementary figures 6 and 7 provide the details of the simulation and the results we obtained. We obtained three main findings: society benefits from POs (and this increases with the size of the bias), the PO is populated by social learners and the share of social learners outside the PO increases with the bias. The first two of these findings are crucial: they show that the results of our baseline model are robust to payoff biased social learning. In what follows, we describe each result in detail.

First, and most importantly, the introduction of the PO increases the fitness of society and this effect is larger if the payoff bias grows. The supplementary figure 6 describes this result for different sizes of the PO (λ) and different biases (ϕ). Several results are noteworthy: i) the larger the bias the higher the fitness; ii) the payoff bias can generate a large increase in the fitness of the PO: while in the baseline model the maximum fitness is 1.15 (standardized with respect to fitness outside the PO equal to 1), when $\phi = 0.9$ the fitness is around 1.3, a substantial increase; iii) larger payoff biases make the size of the PO less relevant for differences in fitness: this is intuitive as larger biases gradually dwarf the benefits of the PO coming from lower social learning costs, iv) the introduction of payoff bias makes that certain sizes of the PO which previously were not beneficial, now they are valuable to society (for example, even $\lambda = 0.9$ is beneficial to society if the payoff bias is large enough).

To understand the mechanics behind the boost on PO fitness, remember that the value of the PO to society depends on the difference between the fitness of social learners outside the PO and inside the PO. Then, notice that the expected level of tuning to the state for social learners outside and inside the PO, q_S and $\widetilde{q_S}$, follow the same recursive equation and therefore in the long run they have the same expected value. Now notice that from the fitness equations above,

$$\widetilde{f_S}(t) - f_S(t) = \widetilde{q_S}(t) - q_S(t) + \beta(t)(c - \tilde{c})$$

taking expectations we have that[2],

$$E[\widetilde{f_S}(t)] - E[f_S(t)] = E[\beta(t)](c - \tilde{c})$$

By looking at the recursive equation for $\beta(t)$ it is possible to observe that for any given $E[\widetilde{f_S}(t)]$, the fraction $\beta(t)$ is increasing in the bias: the second line never occurs (social learners inside the PO will always beat those outside); and the third line is increasing in ϕ, and, for small values of p, it will happen more frequently than the first line, which is decreasing on ϕ (provided that only social learners populate the PO). The fact that $\beta(t)$ is increasing in the bias provides an extra advantage to social learners inside the PO as shown in the equation above.

Second, we find that across all levels of payoff bias, in equilibrium only social learners survive inside the PO. Notice that payoff learning increase the efficacy of social learning because, after the world changes, it makes relatively more likely that social learners will imitate innovators, increasing the speed of transmission of tuned technologies to social learners. As a result, the fitness of social learners increase and so does their share in the population.

[2] In our simulation we corroborate that, in expectation, $\widetilde{q_S}(t) = q_S(t)$.

Third, the level of social (individual) learners outside the PO increases (decreases) with payoff bias. This is displayed in supplementary figure 7 and is readily intuitive: as the bias increases, social learners are the ones which will be mostly benefit as they can get the tuned technology with increasing speed, and thus will increase their fitness and expand.

Summary. By analyzing a model of payoff biased social learning via simulations we find that while payoff bias without a PO doesn't solve Roger's paradox --instead, it simply increases the share of social learners in the population as they become relatively more fit--, introducing a PO increases the fitness of the population and overcomes the paradox. Further, this benefit is increasing in the extent of the bias because as it increases: i) tuned technology can flow faster into the PO (i.e., recovery from a change in environment is quicker) and ii) social learners of the PO, which have a greater fitness, will be increasingly be copied by fellow social learners of the PO, capitalizing more and more on the lower social learning cost. As a consequence, only social learners populate the PO and the size of POs tends to increase. Overall, payoff bias complements the role that POs have for society.

4 Multiple POs, cooperation and group selection

In this section we present two simple models incorporating ideas of group selection, that is, there are more than one PO in the population that compete with one another. These two models are complementary, in the sense that each one focuses and highlights a distinct aspect of the group selection process. In the first model we focus on the consequences of having different firms with different capacity for cooperation (and thus low social learning costs); in the second model, we study how group selection generates the capacity for cooperation in the first place. We provide here a summary of each model and its main finding, and below we analyze each model in detail.

The first model is a straightforward extension of the baseline model in which we assume that the POs are heterogeneous in their capacity to lower the cost of social learning. Achieving a low social learning cost depends on cooperative behavior, particularly in providing consummate effort in teaching and guiding the individual that is copying. We assume that different POs have different capacity in generating conditions conducive to cooperation, rooted in differences such as leadership, contractual devices and/or social norms. The *main finding of the model* is that it shows that POs that have a higher capacity for cooperation, and thus a lower

$\tilde{c}$, will have a larger size than POs with a lower capacity. Thus, we show that group selection will favor the spread of cooperative and social-learning-enhancing POs. The findings of the basic model (i.e., proposition 1 and its implications), also hold in this model with multiple POs.

The second model adapts elements from Bowles et al (2003) and Henrich (2004) and introduces altruism in the process of imitation such that there is a social dilemma: altruism reduces individual fitness but enhances collective fitness. In this model, both individual learners and social learner can be altruistic, where by bearing a cost δ they can reduce the social learning cost that the copier/imitator is experiencing to $\tilde{c}$. Thus, there are four strategies: individual learner, social learner, altruistic individual learner, and altruistic social learner. The fitness of each strategy within a PO is specified as a weighted average of, first, the payoff of the specific strategy with weight $(1 - \mu)$, and second, the average payoff of the strategies inside the PO with weight μ. Replicator dynamics inside the PO operate by contrasting the fitness of the focal strategy to the average of all other strategies across all POs (outside the POs, replicator dynamics operate independently of strategies inside the POs). Thus, unlike the first model, this fitness specification introduces multilevel selection, where the weight μ given to the average fitness of the PO captures the strength of group selection experienced by the PO. The *main finding of the model* is that if μ is small enough, then POs may evolve but they would bring no fitness advantage to society as cooperation would not take hold within POs; in contrast, if μ is large enough, then cooperation takes hold within POs, they become populated only altruistic social learners, and thus a low cost of social learning is attained and POs benefit society.

In sum, these models show that group selection: i) causes the evolution of lower social learning cost within POs, and ii) favors a larger size for POs with lower social learning costs.

4.1 A model of selection among heterogeneous POs

Preliminaries. For learning to occur, assume that in every period each individual gets paired randomly with another individual after which fitness payoffs are generated. The payoffs depend on the learning strategies according to this matrix, which displays the payoff of the person in the rows:

	Individual learner	Social learner
Individual learner	$1 - C$	$1 - C$
Social learner	$1 - c$	$(1 - p)q_{t-1} - c$

This table reflects the payoff of the model in the main body of the paper: An individual learner gets $1-C$ independent on who she gets paired with. A social learner instead, obtains larger benefits from being paired with an individual learner than a social learner as $1 > (1-p)\, q_{t-1}$.

To add a social dilemma to our model, we expand the bottom right quadrant as follows:

	Cooperation	Defection
Cooperation	$(1-p)q_{t-1}-\tilde{c}$	$(1-p)q_{t-1}-c''$
Defection	$(1-p)q_{t-1}-c'$	$(1-p)q_{t-1}-c$

Where $c' < \tilde{c} < c < c''$, and thus social learning entails a prisoner dilemma situation. If both social learners exert consummate effort in the exchange of information, then both benefit from a cost $\tilde{c}$ that is lower than the cost c which occurs when both exert perfunctory effort. However, there is a selfish temptation: if the other party cooperates and exerts effort to teach me, I can be tempted to shirk and avoid putting consummate effort so that I enjoy a cost $c' < \tilde{c}$. I may even put no effort at all, so that the counterparty simply observes and imitates my strategy, without me providing any guidance or teaching at all. Therefore, if I cooperate and the counterparty defects I will have the highest cost of all, c'', as I exert consummate effort in teaching the other party but don't get any teaching in return (I simply imitate without guidance).

In a social dilemma like this, if there is no mechanism to support cooperation; the replicator dynamics will drive the cooperative strategies extinct and the long run equilibrium will feature all defection (Nowak, 2006; Nowak and Rand, 2013). Thus, "selfishness" would prevail. This portray of social learning maps the conception of the evolution of teaching and consummate effort in information sharing as a cooperative act (Laland, 2017; Fogarty et al, 2011).

Thus, in our baseline model, which doesn't specify any mechanism for the evolution of cooperation, the equilibrium will be on the bottom right quadrant. This occurs even if social learners know they could get a lower cost $\tilde{c}$ and a portion of the population is altruistic in nature. Cooperation unravels when part of the population defects and generates a downward spiral into the defection equilibrium (Rand and Nowak, 2013; Boyd and Richerson, 2005; Fehr and Fischbacher, 2003).

One mechanism that allows for the evolution of cooperation is group selection (Henrich, 2004; Henrich, 2016; see Richerson et al, 2016 for a recent review). We leverage this idea to expand our theory to include multiple POs and cooperation, allowing for these to be central aspects of the evolution and role of POs in society.

The basic idea is that there are different POs, each one with a different capacity to generate conditions favorable to cooperation inside their boundaries, and that those POs that excel at it will reach a lower social learning cost and thus will expand more than those that don't. In equilibrium, highly cooperative will have a larger size than POs capable of low levels of cooperation. Notice that this idea is about how group selection harnesses an exogenously given heterogeneity in cooperation capacity across POs; it is not about how group selection can generate this cooperation inside POs in the first place. We address the latter below in section 4.2.

The heterogeneity in cooperation capacity of POs comes from different sources (see the paper by Cordes et al, 2008 for an excellent discussion). Differences may come from how leadership harnesses and crowds in the inherent prosocial and tribal instinct of humans beings (Bowles, 2009; Boyd and Richerson, 2005; Cordes et al, 2008), how social norms of punishment/enforcement (Fehr and Fischbacher, 2003), whether reputational mechanisms are put in place (Nowak and Sigmund, 2005) and even how effective the PO is in using monitoring and formal contractual levers (e.g., team incentives) (Holmstrom and Milgrom, 1991). We are agnostic about how is it that these different drivers affect the maximum level of cooperation a particular PO may attain. All we require is that they generate heterogeneity across POs in the cooperation levels.

Formal analysis. Let's formalize this argument. Assume that there are K different POs where each one display a different $\widetilde{c_k}$. For all POs the cost of social learning satisfies $c' < \widetilde{c_k} < c < c''$. Each PO has its own size λ_k and we assume that the POs will change λ_k in each period by a small amount to improve the average fitness of PO members[3]. POs are secretive, so that only PO members can learn from within their own PO and from the individuals outside any PO (they cannot learn from members of other POs). Likewise, social learners outside the PO only learn from other individuals outside the PO. Social learning is unbiased, that is, models to learn from are selected at random. The recursion equations of tuned population outside the PO is:

$$q(t) = r_I + r_S\,(1-p)q(t-1)$$

And inside the PO k it is the following:

$$\widetilde{q_k}(t) = \widetilde{r_I^k} + \widetilde{r_S^k}\,(1-p)\left[\frac{\lambda_k}{(1-\sum_{i=1,i\neq k}^{i=K}\lambda_i)}\widetilde{q_k}(t-1) + \frac{(1-\sum_{i=k}^{i=K}\lambda_i)}{(1-\sum_{i=1,i\neq k}^{i=K}\lambda_i)}\,q(t-1)\right]$$

[3] We assume that λ_k is inherited without any change from the end of the previous period, and that, at the start of the period, members of the PO make execute a small adjustment of its value by accepting or dismissing members. This adjustment follow a trial-and-error process with learning from the fitness outcome. This assumptions mean that λ_k is not a part of the replicator dynamics.

The expected fitness of all individual learners is $1 - C$, of social learners outside the PO is $f_s = (1-p)q^e - c$, and of social learner inside the PO k is the following:

$$\widetilde{f_S^k} = (1-p)\left[\frac{\lambda_k}{(1-\sum_{i=1,i\neq k}^{i=K}\lambda_i)}\widetilde{q_k^e} + \frac{(1-\sum_{i=1}^{i=K}\lambda_i)}{(1-\sum_{i=1,i\neq k}^{i=K}\lambda_i)}q^e\right] - \left[\frac{\lambda_k}{(1-\sum_{i=1,i\neq k}^{i=K}\lambda_i)}\widetilde{c_k} + \frac{(1-\sum_{i=k}^{i=K}\lambda_i)}{(1-\sum_{i=1,i\neq k}^{i=K}\lambda_i)}c\right]$$

If $K = 1$, then we return to our baseline model of one PO. Assume that $\sum_{i=k}^{i=K}\lambda_i$ is small enough so that there is an internal equilibrium outside the POs (i.e., a positive share of individual and social learners). Therefore, the equilibrium of tuned population outside the PO is equal to $q^e = \frac{r_I}{1-(1-p)r_S}$, and the share of individual learners is $r_I = \frac{p\,[1-(C-c)]}{(1-p)(C-c)}$, both independent of λ_k.

In the PO k, the equilibrium of tuned population is the following:

$$\widetilde{q_k^e} = \frac{\widetilde{r_I^k} + \widetilde{r_S^k}(1-\lambda_k)(1-p)}{1-(1-p)\lambda_k\widetilde{r_S^k}}$$

The analysis of secrecy with a single PO shows that proposition 1 holds, and thus they are populated exclusively by social learners (see proof in section 1 of SI). Given the assumption that $\sum_{i=k}^{i=K}\lambda_i$ is small enough so that there is an internal equilibrium outside the POs, it can be easily shown that proposition 1 also holds for the each of the K POs (given the exclusivity of POs with each other, the proof is actually the same as in the case of a single PO). Thus, given a small enough $\sum_{i=k}^{i=K}\lambda_i$, the PO's fitness is larger than outside and this enhances the overall fitness of society, and, as a corollary, $\widetilde{r_I^k} = 0$ and $\widetilde{r_S^k} = 1$.

PO will calibrate over time their size to maximize the fitness of PO members. The condition that the PO k needs to satisfy to fulfil this goal is,

$$\frac{\partial\widetilde{f_S^k}}{\partial\lambda_k} = \left[(1-p)\frac{\partial\widetilde{q_k^e}}{\partial\lambda_k} - (\widetilde{c_k} - c)\right]\cdot\frac{1}{(1-\sum_{i=1,i\neq k}^{i=K}\lambda_i)} = 0$$

The derivative of tuned population inside the PO on the size of the PO is negative,

$$\frac{\partial\widetilde{q_k^e}}{\partial\lambda_k} = \frac{-q^e p(1-p)}{(1-(1-p)\lambda_k)^2} < 0$$

Thus, exploring $\frac{\partial\widetilde{f_S^k}}{\partial\lambda_k}$ it is easy to see that if $\widetilde{c_k}$ increases (reduces), the PO would need to reduce (increase) λ_k to restore the equilibrium condition. Therefore, a PO with a higher capacity for cooperation would have a larger size in equilibrium. We state this result formally in the following proposition:

Proposition S2: *The POs that display high cooperation and thus a lower $\widetilde{c_k}$, will evolve to be larger than the POs that have low cooperation and a higher $\widetilde{c_k}$.*

A few comments are in order regarding migration and exclusivity.

Migration. In the current formulation, migration is implicit so that in each instance that the PO adjust its size to maximize it member's fitness, some individuals would need to move between POs. If migration is formally model as depending on fitness, we suspect that the results would hold and probably become stronger.

Exclusivity. Notice that exclusivity of POs would still be crucial if group selection is to sustain an outcome that makes social learning useful for society. If a PO reduces its social learning cost via cooperation but it is not exclusive, what will happen is that the PO will expand. In this expansion social learners outside the PO would be gradually replaced by the social learners that populate the PO, until the expected payoff of the PO would be equal to $1-C$. As a consequence, society's fitness would not profit from the cooperation and lower social cost of this particular PO. This analysis suggest that exclusivity would be a PO trait that group selection would need to gradually select as well; here, we are simply assuming it.

4.2 A multilevel model of cooperative social learning within POs

In a large population there are $K \geq 2$ groups composed of $N \geq 3$ individuals that compete against one another. These groups are the POs. We assume that the POs are secretive and thus, social learners outside the POs copy from other members outside the PO randomly (alternatively, this is equivalent to assume that the number of people outside the POs is very large so that in the limit they learn only from outside). Social learners inside a PO copy from agents outside the POs with probability $1-\beta$ and fellow PO members with probability β. This means that while social learners outside copy randomly, social learners inside the PO may have a bias towards PO members.

There are four strategies that can occur inside or outside the PO:

- *Individual learners* with fitness $1-C$.
- *Social learners* select a member of the previous period that they imitate bearing a cost c.
- *Altruistic individual learners.* An altruistic individual learner is willing to sacrifice its own fitness in δ so that when others copy him/her, their cost of copying is reduced to $\tilde{c}$. Thus, as before,

altruism means that the individual learner is willing to dedicate time to teach or mentoring others that copy him. Altruistic individual learners have a fitness $1 - C - \delta$.

- *Altruistic social learners*. An altruistic social learner is willing to sacrifice its own fitness in δ so that when others copy from him/her, their cost of copying is reduced to $\tilde{c}$.

We assume a quasi-live and death replication dynamic, where over long periods of time infinitely lived agents copy the strategy that they observe as having higher fitness. Let q be the tuned population, r_I and r_S be the share of social and individual learning outside POs, r_{AI} and r_{AS} be the shares of altruistic social and individual learning outside POs, and r_A the total share of altruistic types outside the PO. We define analogously $\tilde{q}^{\,i}$, $\tilde{r}^{\,i}_{I}$, $\tilde{r}^{\,i}_{S}$, $\tilde{r}^{\,i}_{AI}$, $\tilde{r}^{\,i}_{AS}$ and $\tilde{r}^{\,i}_{A}$ inside the PO i. The recursions of the tuned population outside the POs and inside the PO i are defined as follows:

$$q(t+1) = (r_I + r_{AI}) + \ (r_S + r_{AS})(1-p)q(t)$$

$$\tilde{q}^{\,i}(t+1) = \left(\tilde{r}^{\,i}_{I} + \tilde{r}^{\,i}_{AI}\right) + \ \left(\tilde{r}^{\,i}_{S} + \tilde{r}^{\,i}_{AS}\right)(1-p)\left[(1-\beta)q(t) + \beta\tilde{q}^{\,i}(t)\right]$$

We define the intermediate concept of a payoff p, which will form part of the general fitness expression we introduce below. The payoffs for each strategy are as follows:

$$p_S = (1-p)E[q] - c + r_A(c - \tilde{c})$$

$$p_{AS} = p_S - \delta$$

$$\tilde{p}^{\,i}_{S} = (1-\beta)(1-p)E[q] + \beta(1-p)E\left[\tilde{q}^{\,i}\right] - c + \beta\tilde{r}^{\,i}_{A}(c - \tilde{c}) + (1-\beta)r_A(c - \tilde{c})$$

$$\tilde{p}^{\,i}_{AS} = \tilde{p}^{\,i}_{S} - \delta$$

$$p_I = \tilde{p}^{\,i}_{I} = 1 - C$$

$$p_{AI} = \tilde{p}^{\,i}_{AI} = 1 - C - \delta$$

For individual learners, being altruistic is always detrimental to their payoff. However, for social learners, we assume that altruism is collectively beneficial but individually detrimental. That is, we assume that $\delta < \beta(c - \tilde{c}) < (c - \tilde{c})$. That is, while the payoff to social learners gets maximized if everyone inside the PO i and/or outside all POs is altruistic, each social learner has the incentive not to be altruistic and save δ. This structure of payoffs captures the social dilemma essence of providing consummate effort in improving social learning.

We assume that the fitness of agents outside POs is equal to their payoff, $f_j = p_j$ with $j \in \{S, AS, I, AI\}$. Let $\overline{p}^{\,i}$ be the average payoff of individuals that belong to PO i, then fitness inside the PO is defined as follows,

$$\widetilde{f}^{\,i}_{j} = (1-\mu) \cdot \widetilde{p}^{\,i}_{j} + \mu \cdot \overline{p}^{\,i}$$

To understand the meaning and significance of μ let's write down the replicator dynamic equation. Let $\dot{r}^i_j$ be the change in the share of strategy j located in PO i, then,

$$\dot{r}^i_j = \widetilde{r}^{\,i}_{j} \cdot \left(\widetilde{f}^{\,i}_{j} - \textstyle\sum_i \sum_j \widetilde{r}^{\,i}_{j} \widetilde{f}^{\,i}_{j}\right) = \widetilde{r}^{\,i}_{j} \cdot \left[(1-\mu) \cdot \left(\widetilde{p}^{\,i}_{j} - \textstyle\sum_i \sum_j \widetilde{r}^{\,i}_{j} \widetilde{p}^{\,i}_{j}\right) + \mu \cdot \left(\overline{p}^{\,i} - \textstyle\sum_i \overline{p}^{\,i}\right)\right]$$

This equation breaks down selection into two forces: i) the first term inside the square brackets captures individual level selection, where a specific strategy within the PO is pitted against the average across all other strategies across all other POs, ii) the second term captures group selection among POs, where the average fitness of the PO is pitted against the average across all POs. This formulation makes clear that μ is the weight that is given to group selection in driving the evolution of different strategies. If $\mu = 0$ then selection will occur only at the individual level, if $\mu = 1$ selection would occur exclusively at the PO level.

We define an ESS equilibrium outside the POs, as the shares p_j so that a small group ε of mutations that enters the population obtains a lower fitness than the existing population. Similarly, the ESS equilibrium inside the PO i is defined as the $\widetilde{r}^{\,i}_{j}$ so that any mutant that replace an existing member of the PO obtains a lower fitness[4].

Given the social dilemma of altruistic strategies, it is easy to see the replicator dynamics will drive the altruistic strategies outside the POs extinct. The same will occur inside the POs if $\mu = 0$. We formalize this with the following proposition:

Proposition S3: *If μ is small enough then in any ESS equilibrium there are no altruists and the average fitness on POs is* $1 - C$.

Proof: Altruists have no advantage outside the POs, so trivially the Equilibrium outside POs correspond to the equilibrium described on our baseline model of the body of the manuscript.

[4] This definition of the ESS inside the PO is needed since POs have a finite number of members and invasions of a certain "measure" ε would make no sense in this context.

What we need to prove is that for any shares of social and individual learners in the POs, any altruistic individual obtains a strictly lower fitness than a non-altruistic mutant if μ is small enough. Therefore, a situation with an altruistic individual does not satisfy the equilibrium condition.

The increase or decrease in fitness experienced by a non-altruistic mutant, either a social or an individual learner, can separated into two effects.

On the one hand, he/she will have a higher fitness by $(1-\mu)\delta$. On the other hand, he/she will affect the average fitness of everyone in the PO –which affects him/her via μ– by decreasing the average by the amount $\frac{\mu\beta}{N}\tilde{r}_S^i(c-\tilde{c})$ and by increasing it by the amount $\frac{\mu\beta}{N}\delta$, where $1/N$ is the mutant share of the PO. Thus, the non altruistic mutant would enjoy a larger fitness if:

$$(1-\mu)\delta + \frac{\mu\beta}{N}\delta > \frac{\mu\beta}{N}\tilde{r}_S^i(c-\tilde{c})$$

Which simplifies to this,

$$(1-\mu)\delta > \frac{\mu\beta}{N}\left[\tilde{r}_S^i(c-\tilde{c})-\delta\right]$$

From where it is easy to see that, if $\tilde{r}_S^i > 0$, and given any value of the remaining parameters, the inequality will hold if μ is made small enough. If $\tilde{r}_S^i$ is equal to 0, then the inequality would hold irrespective of the value of μ.

Given that in this model the PO doesn't bring anything new to the population if μ is made small enough, then the PO will have an internal equilibrium with fitness $1-C$ and shares very similar to those outside the POs (discrete numbers inside the PO, and a the very large size outside of POs, may prevent reaching the exact same share). ■

Thus, the main conclusion of this analysis is that if the competition between POs is weak, then the POs would not evolve their social learning cost advantage, and thus, they would not be able to benefit society.

However, as the reader may already realize, when competition is high, then altruism may thrive within POs. We formalize this in the following proposition,

Proposition S4: *If μ is sufficiently large and $(c-\tilde{c})(1-\beta(1-p)) > p[1-(C+\delta-c)]$, then there exists an equilibrium in which only altruist social learners exist inside the POs and their fitness and payoff is larger than $1-C$. The required level of competition μ for the equilibrium to exist is increasing in the size N of the*

POs and in the cost of altruism (δ) and decreasing in the bias towards learning from the POs (β) and in the advantage generated by altruistic behavior ($c - \widetilde{c}$).

<u>*Proof:*</u> We need to show that if POs are composed of only altruistic social learners, a mutant that replaces an existing member obtains a strictly lower fitness. We analyze first an altruistic individual learner mutant, then an individual learner mutant and finally a social learner mutant.We will show that the mutant obtains lower fitness than other PO members and also members of other POs.

Altruistic individual learner. First notice that, because we assume secretive POs, there is an equilibrium outside the PO, it must be true that $1 - C = (1-p)E[q] - c$, from where we know that $E[q] = \frac{1-C+c}{(1-p)}$. In a PO that is populated exclusively by altruistic social learners, the ratio of tuned individuals inside the PO i is governed by the following difference equation,

$$\widetilde{q}^{\,i}(t+1) = (1-\beta)(1-p)q(t) + \beta(1-p)\widetilde{q}^{\,i}(t)$$

The expected ratio of tuned population in the PO will therefore be

$$E\left[\widetilde{q}^{\,i}\right] = E[q]\frac{(1-\beta)(1-p)}{1-\beta(1-p)} = \frac{(1-\beta)(1-C+c)}{1-\beta(1-p)}$$

This leads to the following fitness expression for the altruistic social learner that populate all the PO's,

$$\widetilde{f}^{\,i}_{AS} = (1-\mu)\cdot\widetilde{p}^{\,i}_{AS} + \mu\cdot\overline{p}^{\,i} = (1-\mu)\cdot\widetilde{p}^{\,i}_{j} + \mu\cdot\widetilde{p}^{\,i}_{AS} = \widetilde{p}^{\,i}_{AS}$$

$$\widetilde{f}^{\,i}_{AS} = (1-p)\left[(1-\beta)E[q] + \beta E\left[\widetilde{q}^{\,i}\right]\right] - c + \beta(c-\widetilde{c}) - \delta$$

For $\widetilde{f}^{\,i}_{AS}$ to be larger than the fitness of an altruistic individual learner $1 - C - \delta$, then it is required that $(c-\widetilde{c})(1-\beta(1-p)) > p[1-(C-c)]$. This condition shows that the right hand side benefits of having an individual learner (i.e., getting the tuned technology in periods where nature changes but at a higher cost), must be lower than the cost that it generates by not copying other PO members and thus, by not generating the advantage $(c-\widetilde{c})$ for the fraction $1-\beta(1-p)$ that an altruistic social learner would generate[5].

Observe that the condition $(c-\widetilde{c})(1-\beta(1-p)) > p[1-(C-c)]$ is sufficient but not necessary when comparing with another member. We compared the fitness of an altruistic social learner inside the PO when all PO members are altruistic social learners versus the fitness of an altruistic individual learner. Because

[5] This fraction means the individual learner does provide the benefit to anyone that copies him/her but that he/she doesn't generate the benefit to the β share that a social learner would generate by copying others in times of environmental stability.

N in the PO can be small, the precise comparison would be to use the fitness of an altruistic social learner inside the PO when one member of the PO is an altruistic individual learner and all others are altruistic social learners. Mathematically this is the correct condition to study, $(1-\mu)\cdot\tilde{p}^{\,i}_{AS}+\mu\cdot\bar{p}^{\,i}_{j}>(1-\mu)\cdot\tilde{p}^{\,i}_{AI}+\mu\cdot\bar{p}^{\,i}_{j}$, which simplifies to $\tilde{p}^{\,i}_{AS}>\tilde{p}^{\,i}_{AI}$. This is the condition we assessed but $\tilde{p}^{\,i}_{AS}$ was not calculated assuming that the PO has one individual learner. It is easy to see that $\tilde{p}^{\,i}_{AS}$ would only grow if we had considered the individual learner mutant, and thus, the inequality will always hold with slack when the condition holds.

The argument in the previous paragraph makes it clear that the mutant obtains a lower fitness than the members of other POs as long as the condition holds.

Individual learner mutant. This situation is the same as before only that the individual learner doesn't pay the cost of altruism. Thus, by comparing $\tilde{f}^{\,i}_{AS}$ to $1-C$ we find that the condition is now a bit more restrictive: $(c-\tilde{c})(1-\beta(1-p))>p[1-(C+\delta-c)]$.

Exactly for the same reason we explained in the note above the condition is necessary and sufficient to show that the mutant also obtains a lower fitness than members of other POs.

Social learner mutant. Imagine that one individual learner mutant replaces one altruistic social learner. The incremental fitness of the mutant relative to the altruistic social learner has two elements: i) because she doesn't put effort when others are learning from her, her fitness increases by $\delta(1-\mu)$, ii) the average fitness of the PO decreases by $\frac{\beta\mu}{N}(c-\tilde{c})$ and increases by $\frac{\beta\mu}{N}\delta$. Therefore the PO can't be invaded by individual learners if the following condition holds (this is the reverse of condition described in the proof of proposition S3 above),

$$(1-\mu)\delta<\frac{\mu\beta}{N}[(c-\tilde{c})-\delta]$$

Rearranging,

$$\mu>\frac{\delta N}{\delta N+\beta[(c-\tilde{c})-\delta]}$$

Thus, the level of competition μ must be higher than a threshold in order to fend off a social learner mutant. This threshold is increasing in the size of the PO (N) –as one would expect to to the problem of free-riding in larger groups– and in the cost of altruism (δ) and decreasing in the bias towards learning from the PO (β) and in the advantage generated by altruistic behavior ($c-\tilde{c}$).

In this case because we are only considering social learners if the mutant obtains a lower fitness than her own PO members she will also obtain a lower fitness tan the members of other POs that only have altruistic social learners. ■

Although it is significantly more onerous it can be shown that if μ is sufficiently large then the only ESS equilibrium has only altruist social learners in the POs.

Summary. The results of our baseline model hold when one does not simply assume cooperation within POs but instead model its emergence as an outcome of group selection among many POs. Thus, the lower social learning cost within POs is now endogenous, not an assumption of the model. The level of group selection required for cooperative POs to evolve is higher when altruism is more difficult (due to a larger POs size or a larger cost of altruism) but lower when altruism is more beneficial (due to a higher impact $c - \tilde{c}$ or a better utilization of its benefits, captured by β).

5 Incremental innovation within POs

The interpretation of the model can be extended by acknowledging that in order to adopt and use/apply a technology, both learning and production are required. One has to learn not only what to do -- for example, learn that net fishing generates a good catch in the current environment -- *but also develop an understanding and capacity of how to do it* -- for example, know which are the best raw materials to produce the net, how to assembly it properly, how to use it best, when and where it provides more benefits, and so on. An individual learner has to figure out these two challenges, learning what to do and how to do it; similarly, a social learner has to copy both elements. Current models simply compress these two dimensions into the single act of copying a technology.

By adding the production stage, we can introduce the idea of incremental innovation, defined as the process through which the execution of an activity can be improved bit by bit over time. Incremental innovation contrasts with the concept of radical or disruptive innovation which overhauls the current technological standards providing a novel solution to a problem; this typically renders the previous technology obsolete, or at least, leaves it at a great disadvantage (Henderson, 1993). And while incremental innovation can be a part of social learning that includes a production/refinement stage, radical innovation can be associated with individual learning: it requires that individuals spend considerable effort to understand the nature of a problem/landscape and then generate a novel and well-adapted technology. Notice that as part of this "enhanced" notion of social learning, incremental innovations are not protected from changes in the environment: no matter how refined the technique might become over time, the environment can still change in favor of a different technology. In simple, if $p = 1$, the old technology is no longer useful.

By allowing agents to engage both in learning and production, the parameters C, c and $\tilde{c}$ would now include not only learning costs but also the costs of production. Thus, our theoretical model can be extended by assuming that in every instance of copying, the social learner can introduce incremental improvements, which can then be passed down to other social learners (which would further add incremental improvements, and so on). This process can make the difference between C and c to be large, as it will capture not only that "discovering" is harder than "copying" but also that through copying, improvements are generated.

Now, it is known that POs have a natural capability to generate incremental innovations. This fact has been extendedly documented (Henderson, 1993; Arora et al, 2018): firms facilitate incremental innovations but struggle with radical innovations --which are, instead, typically introduced by outsiders. This would mean that this process of incremental innovation, which thrives within POs, would be reflected in an ever larger the difference between c and $\tilde{c}$. The organizational learning literature in economics and management shows indeed that these enhancements to production techniques, and the ensuing reduction in costs, can be very large (Argote and Miron-Spektor, 2011; Levitt et al., 2013). As discussed in this literature, organizations provide a unique environment for social learning and gradual enhancement of technologies/activities, such as copying fidelity, easier communication, enhanced codification, informal networks, stability of working relations, etc. etc.

By introducing this interpretation to our model, one could argue that $\tilde{c}$ goes down gradually over time, becoming much smaller than c. What this does, is that it allows for a much larger impact of POs on society. Now the PO has an innovation role: by adding small improvements over time it can provide a very large benefit to society.

Notice that in this appendix we simply modify the interpretation of the cost parameters; a fuller treatment and development of the idea of production, and the distinction between incremental and radical innovation, is beyond the scope of this paper and would require in a separate manuscript.

6 Details of the empirical analysis

Descriptive statistics

In supplementary figure 10 we map the 173 societies used in table 1 of the manuscript. In supplementary tables 1 to 5 we provide several descriptive statistics of this sample.

Endogeneity

Regarding *omitted variable bias*, in supplementary table 6 we replicate column 2 of table 1 and columns 2, 5, and 8 of supplementary table 10 (in this table we use alternative dependent variables; see section "Robustness checks" below). Assuming a maximum R-square 0.95 --there is an inherent measurement error of ethnographic data-- the "delta" is on average 0.97. This means that selection on unobservables would need to be at least 0.97 times the selection on observables in order to overthrow our results. A delta of 1 suggest a low threat of omitted variables, particularly if a comprehensive set of controls is used (Oster, 2016).

Regarding *reverse causality*, this threat can be present both in the presence of activities and in the use of POs. We analyze each one in turn.

An important proposition of cultural evolution is that a larger and more interconnected population would generate more cumulative culture (Henrich, 2016). In our case, this would translate into a higher presence of activities in the society, which could be channeled through the interaction with POs. This could generate an upwards bias in our baseline estimation of the impact of POs. To address this issue, we instrumented "% of presence" using two variables: "sex differentiation" (which we detail in section 4.3.2) and the index of "kinship tightness" developed by Enke (2019). For the first instrument, there is evidence that the presence of activities coevolved with sex differentiation in premodern societies (Haun and Over, 2013). For example, men specialize in large game hunting, while women specialize in gathering. Sex differentiation might affect population size if it affects fertility. We contend that conditional on the total amount of activity executed by women, and therefore controlling for the time restriction that differentiation might place on female fertility, the exclusion restriction should hold.

Kinship systems regulate the pattern of relatedness in society through family structure (e.g., independent nuclear families vs. extended families), marriage patterns (e.g., cousin marriage allowed vs. forbidden), and descent (e.g., unilineal vs bilateral descent group). Kinship tightness is a key variable affecting social organization of a society (Enke, 2019). Tight kinship generates high in-group bias, less cooperation with outsiders, strong conformism, and local institutions. The opposite occurs with loose kinship, with the consequence of being much more open to external groups. We argue that kinship tightness affects positively the presence of activities in society. A society with loose kinship, and therefore more "open", would be more inclined to source part of the basic activities from neighbors; in contrast, a society with tight kinship, and therefore more "closed", would have no alternative but to source some of the basic activities internally. The exclusion restriction for "kinship tightness" is sustained on the documented ancestral origins of kinship systems: its origin can be traced to the societies from which the focal society descends from (Passmore and Jordan, 2017). Thus, exogeneity is plausible, particularly when controlling for agriculture intensity and settlement type[6]. Further there are no a priori reasons relating kinship tightness to larger or smaller populations through changes in fertility, the main channel that kinship might affects[7].

In the supplementary table 7 on the online appendix, we present the instrumental variables estimations. We find that both sex differentiation and kinship tightness are positively and significantly related to the presence of technologies in the first stage, and that they are not weak. The Hansen-test indicate that the instruments are exogenous, in line with the theoretical arguments laid out above. In the second stage, the results do not change from those of the table 1: the presence of technologies increases the local community size, but only when PO are in place. The IV estimation is robust is we use total population of the society as the dependent variable (see section "Robustness checks" below).

Reverse causality could also affect the variable "% within PO". If there is a minimum size for POs and/or specialization is favored by the extent of the market, large populations might make it easier to have POs. We address this issue by instrumenting the presence of POs following the idea of Depetris-Chauvin and Ozak (2017)[8]. These authors explore the extent to which genetic diversity in the population of the societies in the ethnographic atlas is related to the presence of POs. The argument is that a genetically diverse population has

[6] There is evidence that kinship tightness evolved to match the needs of agricultural subsistence, away from nomadism (see Enke, 2019).

[7] For example, the polygamy-fertility literature is not at all conclusive. Accordingly, and consistently with Enke (2019), we do not find a relationship between kinship tightness and population in our data, conditional on covariates.

[8] An additional way to study the problem of reverse causality of "% within PO" is to check whether its coefficients change with the level of the dependent variable: reverse causality would predict that the positive relationship of the coefficients would be much stronger at higher levels of population. The result displayed in the panel D of figure 2 above argues against this: there we show that the impact of PO is exerted throughout the different size categories of the dependent variable "size of local community". For the case of the dependent variable "total population" we replicated column 2 of table S8 using a quantile regression. In panel A of figure S11 of this SI we display the value of the coefficient related to "% presence" as it varies across the dependent variable; on panel B of figure S11, we do the same for the interaction term "% presence x % within PO". In both cases, the graphs show that the positive impact of POs on population is exerted evenly across different population sizes.

many different skills in place, which would lead to the creation of different specialized groups. These authors instrument genetic diversity using the distance of the society from East Africa (specifically, modern day Ethiopia), which is the origin of the spread of the human species out of Africa (starting approximately 80,000 years ago). As the distance from Africa increases, the diversity within a society goes down, a phenomena known as serial founder effect (Ramachandran et al., 2005). The authors find substantial evidence in favor of their arguments: distance reduces genetic diversity, which in turn decreases the presence of POs.

In supplementary table 8, we follow these authors and use "Distance from Africa" as an instrument for "% within PO". We measure the distance from Addis Adaba in east Africa to the focal society; for societies in America, we calculate the distance going through the Bering strait. We do not use the mediating variable of genetic diversity, and thus, we implement the "reduced form" model of Depetris-Chauvin and Ozak (2017)[9]. In column 1, we present the first stage. Consistent with Depetris-Chauvin and Ozak (2017), the distance from Africa reduces the presence of POs in societies. Although the Cragg-Donald test (reported in columns 2 and 3) indicates that the instrument is relevant, by comparing the values with Stock and Yogo (2002), we cannot rule out weakness in the instruments. To address this issue, in the second stage of columns 2 and 3, we use the limited information maximum likelihood (LIML) technique, which partially mitigates the problem of weak instruments. The results that we obtain with both dependent variables are consistent and supportive of our predictions. By comparing the values of the coefficient with those of table 1 and supplementary table 8 of the online appendix, we also find that the coefficients exhibit increases in their size.[10]

Comparative statics

We test the comparative statics derived from our model, which are summarized in figure 2 of the main body. The econometric model that we use is the following:

$$Population_i = b1 + b2 \ x \ \%Presence_i + b3 \ x \ \%Presence_i \ x \ \%withinPO_i + b4 \ x \ \%Presence_i \ x \ Z \\ + b5 \ x \ \%Presence_i \ x \ \%withinPO_i \ x \ Z + Controls_i + Error_i \quad (8)$$

[9] Therefore, our results cannot be interpreted as finding any causal effect of genetic diversity on any economic variable in the society. Our reduced form allows other mechanisms to impact the presence of POs. For example, given the non-randomness of the migratory sampling process, it could also be the case that traits are lost. As migrant groups are typically small, the likelihood of loss increases due to drift.

[10] We also instrumented "% presence" and "% within PO" at the same time. We used the three instruments simultaneously following Wooldridge (2010, chapter 8). The results are consistent with table 3 and table 4. The coefficient for "% presence" is 1.32, and that for "% within PO" is 3.83. However, statistical significance is lost. The instruments retain their properties: for strength, they surpass the Stock and Yogo thresholds on strength; for exclusion, the Hansen test indicates that the instruments are valid.

In this model, we generate a triple interaction to explore whether the impact of POs is affected by the variable *Z*. We use different variables as *Z* in order to proxy for the different parameters *p*, *c* and $\widetilde{c}$, in addition to the prevalence of secrecy. If the coefficient *b5* is positive (negative), then the impact of PO is enhanced (diminished) by the variable *Z*. Of course, the mapping between our proxies and the parameters of the model is not perfect. However, we believe that if several imperfect proxies confirm the comparative statics of the model, confidence can be gained. The variables we study are the following:

Uncertainty. To test the impact of environmental uncertainty, the parameter p of the model, we use climate unpredictability. Climate data have already been successfully used to empirically test the parameter p in the baseline cultural evolution model (Giuliano and Nunn, 2019). The D-PLACE dataset follows the methodology of Colwell (1974) and reports "temperature unpredictability" and "precipitation unpredictability" using yearly data between 1901 and 1950, the period that has the largest proportion of ethnographies in the ethnographic atlas (see supplementary table 5). We multiply these two measures to obtain our measure of climate unpredictability. Consistent with the comparative statics of the model, the results from the column 1 of supplementary table 9 show that the impact of POs on population decreases when climate unpredictability is high. A joint t-test shows that the impact of PO is again highly significant and, importantly, moderated by climate unpredictability. This result is shown in the panel A of supplementary figure 12.

Social learning costs. We studied three variables that decrease the costs of social learning. First, the SCCS provides information about how prevalent apprenticeship and teaching are in the society. "Apprenticeship" is a dummy variable that we computed from the variables v427 and v428 of the SCCS that measure the extent of guidance and/or formal schooling in late boys and girls, respectively. The dummy takes the value of 1 when one or both of these variables indicate that the society displays "predominant apprenticeship", or when "formal schooling is frequent and typical", and zero otherwise. Clearly, if schooling and apprenticeship are predominant in society, this will decrease both c and $\widetilde{c}$. Given that a lower c has an ambiguous impact on the fitness of POs (see panel A of figure 2), but a lower $\widetilde{c}$ has an unequivocal increase in the fitness of POs (panel B of figure 2), we predict that "apprenticeship" should boost the impact of POs on population size[11]. The results are presented in column 2 of supplementary table 9 and are consistent with our prediction: the positive impact of POs on population is stronger if apprenticeship is predominant. We graph this result in panel B of supplementary figure 12. The second variable that reduces cost of social learning is "sex differentiation". The variables v44 to v54 of the EA provide information about the extent to which the eleven activities are executed by women and/or men. For each activity we coded a dummy that took the value of 1 in case the activity was executed by "males only or almost alone" or by "female only or almost alone". Then, we added these dummies and divided the result by the total number of activities that have available information. We label this variable "sex differentiation", and it captures the percentage of activities that are performed by either sex exclusively. There is plenty of evidence that social learning is facilitated by similarity, in which sex plays an important part (Henrich, 2016, p. 44; Fairlie et al., 2014; Losin et al., 2012). Similar to "apprenticeship", this variable reduces both c and c, and therefore, our model predicts that more sex differentiation would lead to an increase in the impact of POs. This is what we find in the third column of supplementary table 9: we obtain a positive and significant coefficient for the interaction term. The third variable that reduces the cost of social learning is "Trust". We use the variable v335 of the SCCS, which measures the degree to which trust in inculcated in childhood in the society. This variable is ordinal, with 0 meaning "no inculcation or opposite trait" and 9 meaning "extremely strong inculcation". As before, high "trust" decreases both c and $\tilde{c}$, leading to the prediction of a higher impact of POs. The result, displayed in column 4 of supplementary table 9, is consistent: POs have a larger positive impact on population when trust is high. This result is shown in panel C of supplementary figure 12.

Secrecy. Finally, we analyze the impact of the variables "Honesty" and "Generosity". These are the variables v336 and v334 of the SCCS and are analogous to v335, namely, a categorical variable identifying inculcation of honesty and generosity in childhood. It is possible to identify these variable with a decrease in the degree of secrecy in the POs and therefore a boost in their fitness. Oftentimes, secrecy is related to selfish behavior, a desire to keep useful knowledge proprietary. The zeal to maintain secrecy could also benefit from dishonest behavior, deflecting requests to share knowledge with lies. It would also be possible to relate "honesty" and "generosity" to a decrease in the costs of social learning. When agents are generous and honest,

[11] It could also be argued that this dummy would be more tightly connected to a decrease in c than one in c because teaching and apprenticeship probably coevolved with POs. This possibility would reinforce our prediction.

it is very likely that communication and learning would improve. In both cases, the prediction from our model is clear: these variables should increase the fitness benefits of POs. The results are presented in columns 5 and 6 of supplementary table 9 and are consistent with the prediction of our model. In the panel D of supplementary figure 12, we display the results for generosity.

Robustness checks

Robustness to alternative dependent variables. We use three alternative dependent variables: "total population", "population density", and "cultural complexity". The variable "total population" is obtained from the Ethnographic Atlas and is a continuous variable that indicates the total population of the society. We use natural logarithms to normalize its distribution. In the columns 1, 2 and 3 of supplementary table 10 we display the results. The results do not change: total population increases with the presence of technologies, but only when these are executed by POs; and the positive impact of POs on total population increases when climate unpredictability is low.

"Population density" is a categorical variable measuring persons per square mile (see supplementary table 10 for details). The results presented in columns 4, 5 and 6 of supplementary table 10 show that our findings are also robust to using this alternative dependent variable.

Finally, the Ethnographic Atlas reports "cultural complexity", a variable that tries to capture the degree of cultural sophistication (see supplementary table 10 for details). Overall, the results reported in columns 7, 8 and 9 of supplementary table 10 show that the results are robust to this alternative dependent variable.

Alternative explanations. There are three main alternative explanations for our empirical results. We address each one in turn. First, it could be argued that the positive impact of POs is due to POs generating improvements in terms of the costs of individual learning, rather than the costs of social learning. To assess this possibility, we studied a model in which a PO decreases C instead of c. A model with this characteristic generates POs populated entirely by individual learners, and importantly, their benefits are increasing in the uncertainty parameter p. The latter implication is directly contradicted by our interaction with uncertainty: we find that uncertainty reduces the impact of POs on population (column 1 of supplementary table 9 and panel A of supplementary figure 12). The former implication -- POs are populated by individual learners -- is rebutted by simple perusal of organizational reality: in general, social learners dominate individual learners inside POs (and the opposite occurs in the market).

A second alternative explanation is related to trade. Specialized POs might have a positive impact on population because they are a marker for the presence of trade in the society, and trade could be the fundamental driver of larger populations and the key force behind the evolution of specialized POs. To test this alternative explanation, we use as controls three variables from the SCCS that proxy for the presence of trade in the society (see the supplementary table 11 for details). Across these three controls, the coefficient of interest decreases in size by an average of 16% but remain statistically significant. These reductions in the size of the coefficients indicate that some of the impact of POs is indeed generated through trade benefits, but that it is not the main mechanism. Instead, this result is consistent with our proposition 3, which states the origin of specialized POs is driven by the need to make social learning useful in society, without requiring trade as a force for its evolution.

A second way to assess the alternative explanation of trade is by exploring the interaction with uncertainty. The literature on trade has proposed and documented that trade (and thus the specialization that it drives) is particularly useful to mitigate the effects of shocks to local productivity, such as weather changes (Burgess and Donaldson, 2010). Therefore, if the benefits of trade are the key driver of the impact of POs, we should find a positive interaction of POs with uncertainty; however, we find the opposite in our results (column 1 of supplementary table 9 and panel A of supplementary figure 12).

The third alternative explanation for the origin of POs is that they emerge as a result of political complexity. The idea is that having a complex political organization in the society (e.g., centralized state) allows the society to better define, monitor and enforce POs. Thus, it might be that political complexity is really driving both the presence of POs and a larger population. In supplementary table 12, we add the variable "political hierarchy" as a control. This variable measures the levels of political authority: from "no political authority beyond the local community" to "four levels of authority (e.g. large states)" (see the table for details on the variable). The results shows that the coefficients are reduced by 23% on average across the different dependent variables we use. This indicates that the alternative explanation carries some weight, but not sufficient to overthrow our results[12].

Data stringency. The results are also robust to being less restrictive regarding the available information about the activities (supplementary table 13). In many societies, there is information about only a portion of the

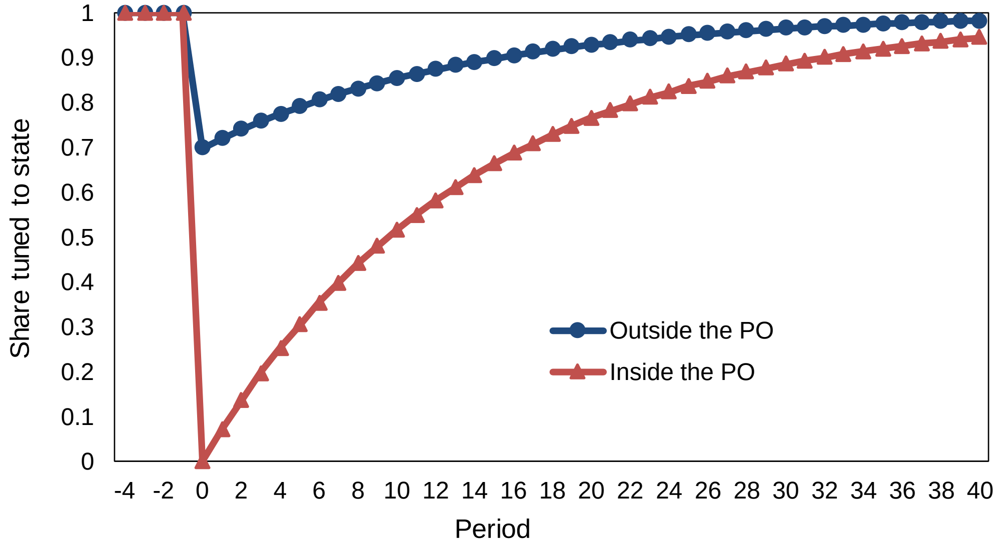
[12] It could also be argued that political complexity is driven by POs in the first place. If that is the case, including this control would be biasing downward the true impact of POs.

activities (e.g., inland societies do not have boat building). We changed the minimum number of activities that have available information in the society, and the results do not change. Further, in columns 5 and 6 of supplementary table 13, we restricted the sample to regions that have at least 2 and 3 societies in them, leading to a loss of 7 and 19 societies, respectively. The results remained unchanged.

7 Figures and tables

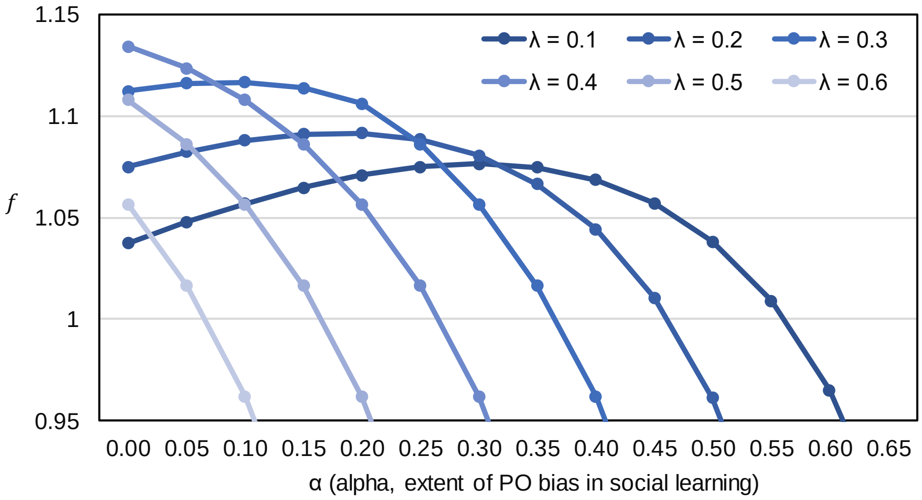
Supplementary Figure 1. Adaptation to a change in environment inside and outside the PO.

Notes: To build this graph we assume one generation in the population. We use C=0.6, c=0.45, $\tilde{c}$=0.3, and p=0.1 to compute the equilibrium level of r_S, $\tilde{r}_S$, r_S and $\tilde{r}_S$. Then, we use equations (1) and (2) to graphs the dynamics of q and $\tilde{q}$. We assume that in period 0 there is an environmental change (p = 1), so the rest of the periods have p = 0.



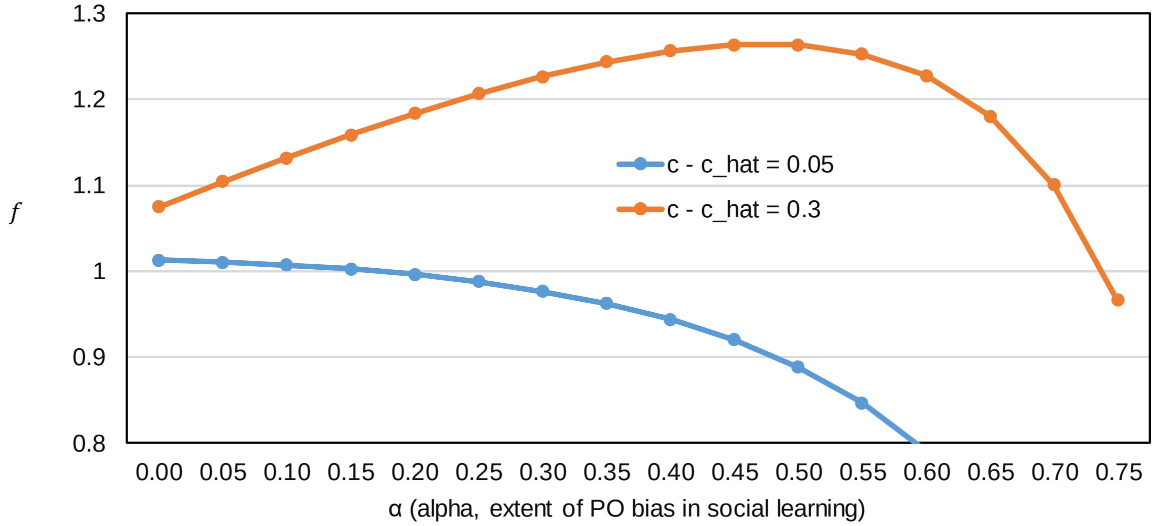
Supplementary Figure 2. Impact of PO-biased social learning on the fitness of the PO.

Notes: We use C=0.6, c=0.45, $\tilde{c}$=0.3, and p=0.1 (the fitness values are multiplied by 2.5 as in the figure 1 and 2 of the manuscript in order to standardize with respect to individual learners).



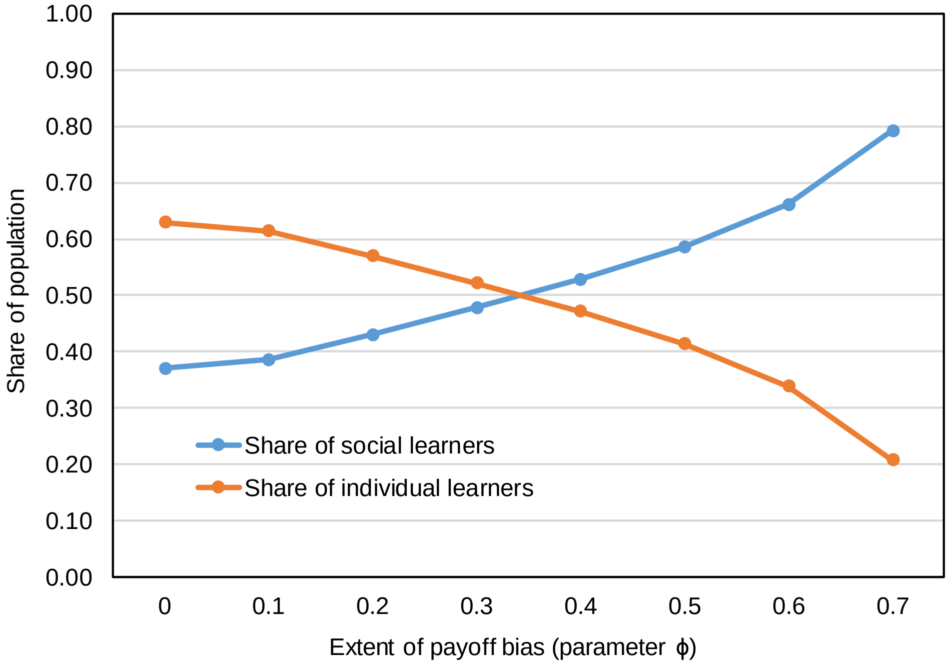
Supplementary Figure 3. Impact of PO-biased social learning on the fitness of the PO varying the advantage in social learning inside the PO.

Notes: In the blue line we use λ=0.1, C=0.6, c=0.45, $\tilde{c}$=0.4, and p=0.1; In the orange line we change $\tilde{c}$ to be equal to 0.3 (the fitness values are multiplied by 2.5 as in the figure 1 and 2 of the manuscript in order to standardize with respect to individual learners).



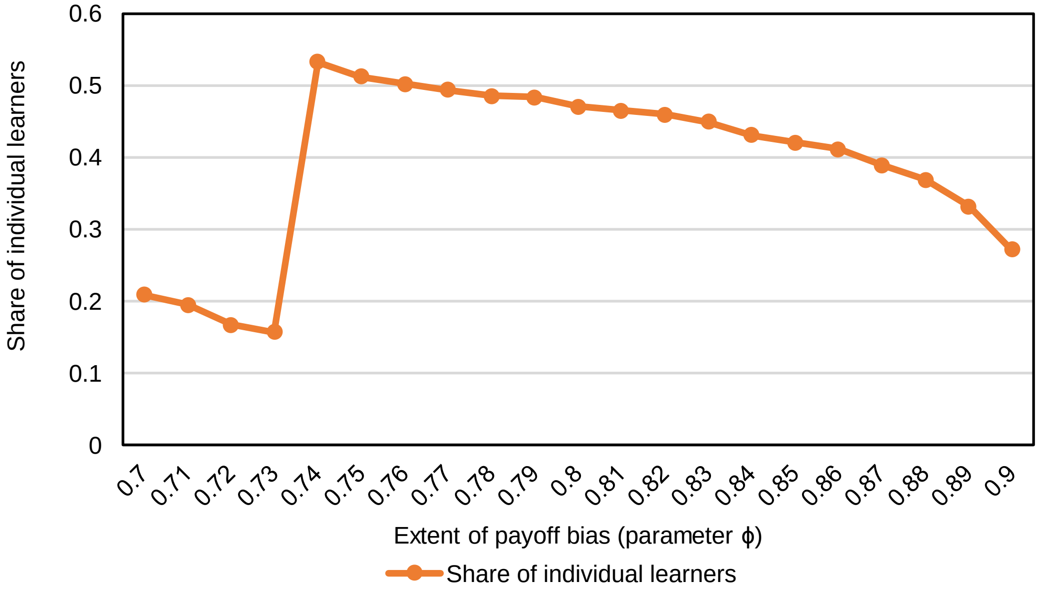
Supplementary Figure 4. Share of social and individual learning depends on the extent of payoff bias.

Notes: We use C=0.6, c=0.45 and p=0.1. For this simulation we used the following seeds in period 0: r_I=0.5; q_S=0.7. For each bias ϕ, we iterated over 150.000 periods and report the average between the periods 50.001 and 150.000. The share of each strategy changes according to the following replicator dynamic: change in share = 0.01 x (share in t-1) x {(fitness of strategy in t-1) - (average fitness in t-1)}.



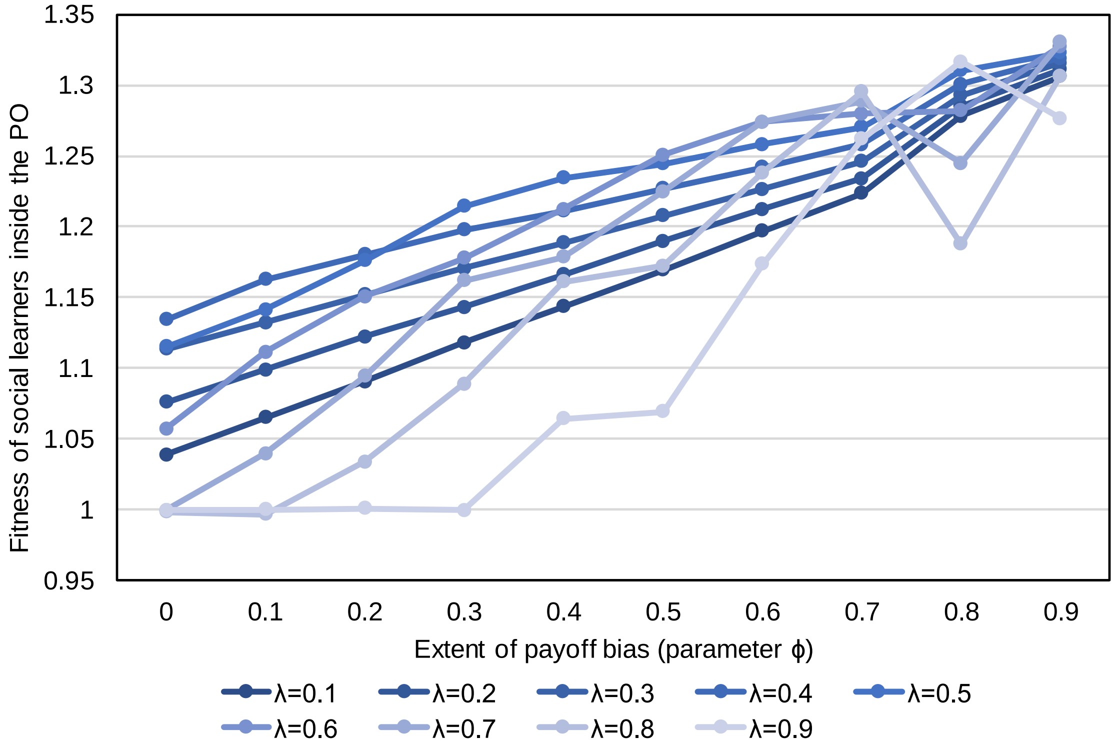
Supplementary Figure 5. Share of individual learning and social learner fitness for high level of payoff bias.

Notes: We use C=0.6, c=0.45, $\tilde{c}$=0.3 and p=0.1 (the fitness values are multiplied by 2.5 as in the figure 1 and 2 of the manuscript in order to standardize with respect to individual learners). For this simulation we used the following seeds in period 0: r_I=0.5; q_s=0.7; $\tilde{r_I}$=0.5; $\widetilde{q_s}$=0.7. For each bias ϕ, we iterated over 150.000 periods and report the average between the periods 50.001 and 150.000. The share of each strategy outside the PO changes according to the following replicator dynamic: change in share = 0.01 x (share in t-1) x {(fitness of strategy in t-1) - (average fitness in t-1)}. The same replicator dynamic governs the changes of the two strategies inside de PO (i.e., the replicator dynamics are independent).



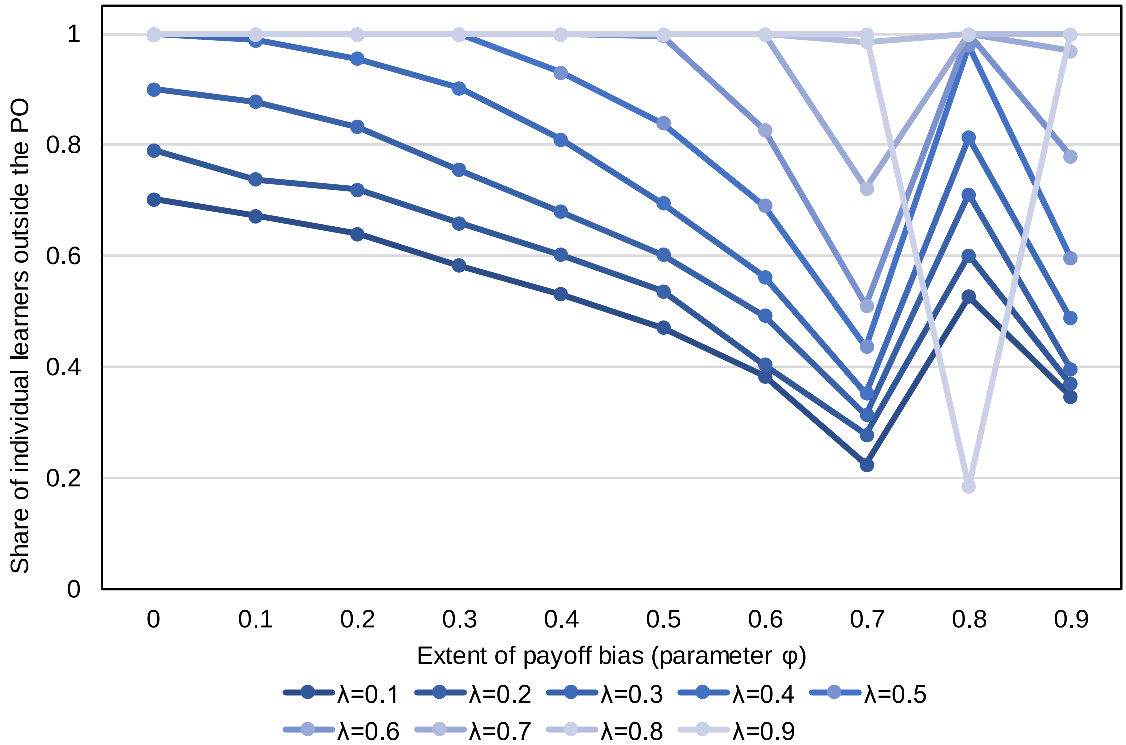
Supplementary Figure 6. Fitness of social learners inside the PO for different payoff biases.

Notes: We use C=0.6, c=0.45, $\tilde{c}$=0.3 and p=0.1 (the fitness values are multiplied by 2.5 as in the figure 1 and 2 of the manuscript in order to standardize with respect to individual learners). For this simulation we used the following seeds in period 0: r_I=0.5; q_s=0.7; $\widetilde{r_I}$=0.5; $\widetilde{q_s}$=0.7. For each bias ϕ, we iterated over 150.000 periods and report the average between the periods 50.001 and 150.000. The share of each strategy changes according to the following replicator dynamic: change in share = 0.01 x (share in t-1) x {(fitness of strategy in t-1) - (average fitness in t-1)}.



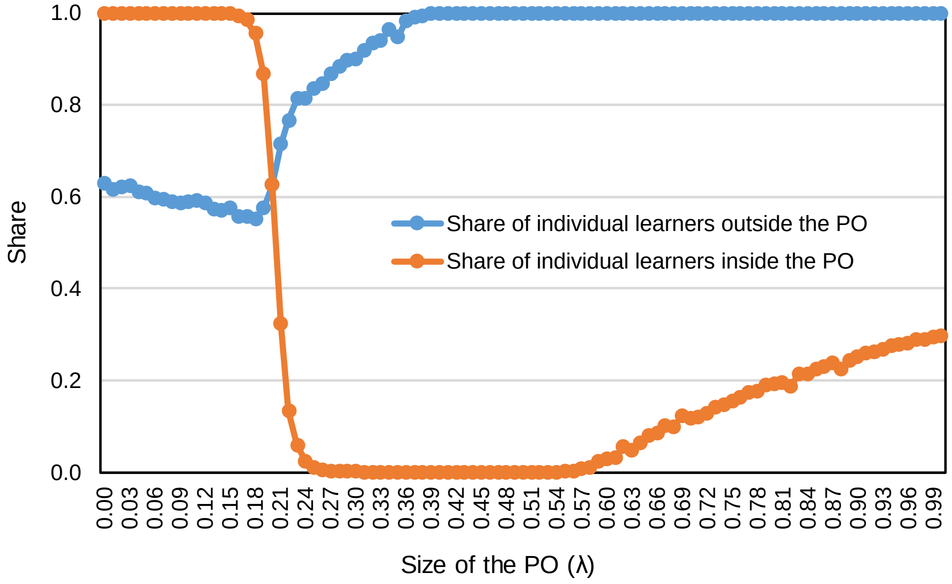
Supplementary Figure 7. Share of individual learners outside the PO for different payoff biases.

Notes: We use C=0.6, c=0.45, $\tilde{c}$=0.3 and p=0.1. For this simulation we used the following seeds in period 0: r_I=0.5; q_s=0.7; $\widetilde{r_I}$=0.5; $\widetilde{q_s}$=0.7. For each bias ϕ, we iterated over 150.000 periods and report the average between the periods 50.001 and 150.000. The share of each strategy changes according to the following replicator dynamic: change in share = 0.01 x (share in t-1) x {(fitness of strategy in t-1) - (average fitness in t-1)}.



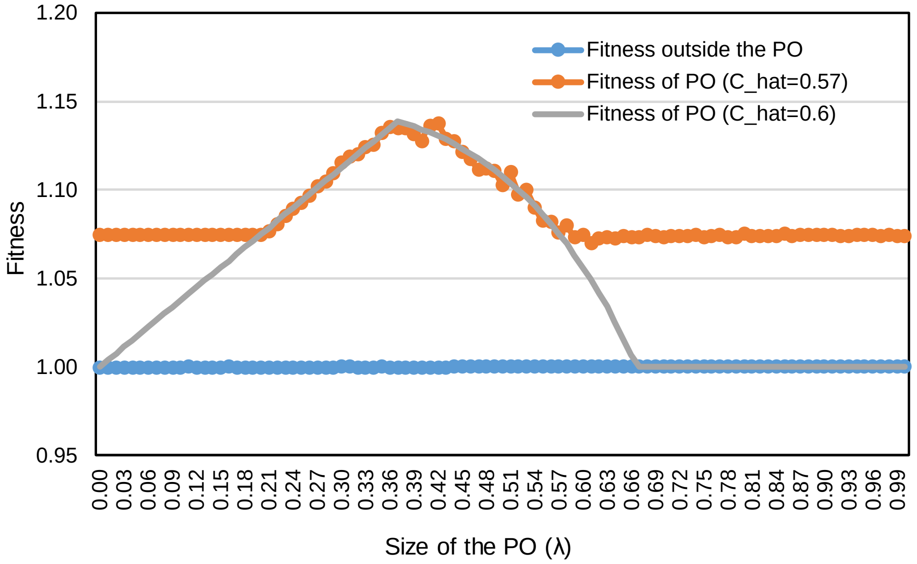
Supplementary Figure 8. Share of individual learners inside and outside the PO.

Notes: We use C=0.6, $\widetilde{C}$ =0.57, c=0.45, $\widetilde{c}$=0.3 and p=0.1 (the fitness values are multiplied by 2.5 as in the figure 1 and 2 of the manuscript in order to standardize with respect to individual learners). For this simulation we used the following seeds in period 0: r_I=0.5; q_s=0.7; $\widetilde{r_I}$=0.5; $\widetilde{q_s}$=0.7. We modified lamda by 0.01 at the time. For each lamda, we iterated over 150.000 periods and report the average between the periods 50.001 and 150.000 as the long run equilibrium. The share of each strategy changes according to the following replicator dynamic: change in share = 0.01 x (share in t-1) x {(fitness of strategy in t-1) - (average fitness in t-1)}.



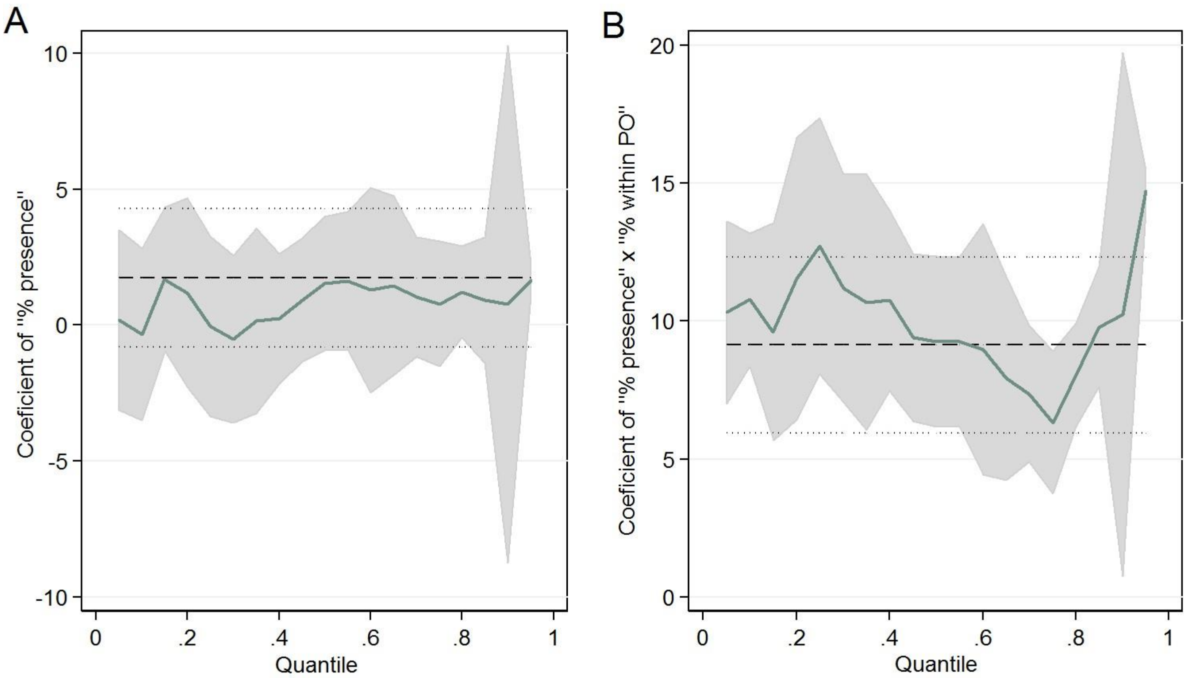
Supplementary Figure 9. Fitness of the PO when the PO improves both social and individual learning cost.

Notes: We use C=0.6, $\widetilde{C}$ =0.57, c=0.45, $\widetilde{c}$=0.3 and p=0.1 (the fitness values are multiplied by 2.5 as in the figure 1 and 2 of the manuscript in order to standardize with respect to individual learners). For this simulation we used the following seeds in period 0: r_I=0.5; q_s=0.7; $\widetilde{r_I}$=0.5; $\widetilde{q_s}$=0.7. We modified lamda by 0.01 at the time. For each lamda, we iterated over 150.000 periods and report the average between the periods 50.001 and 150.000 as the long run equilibrium. The share of each strategy changes according to the following replicator dynamic: change in share = 0.01 x (share in t-1) x {(fitness of strategy in t-1) - (average fitness in t-1)}.



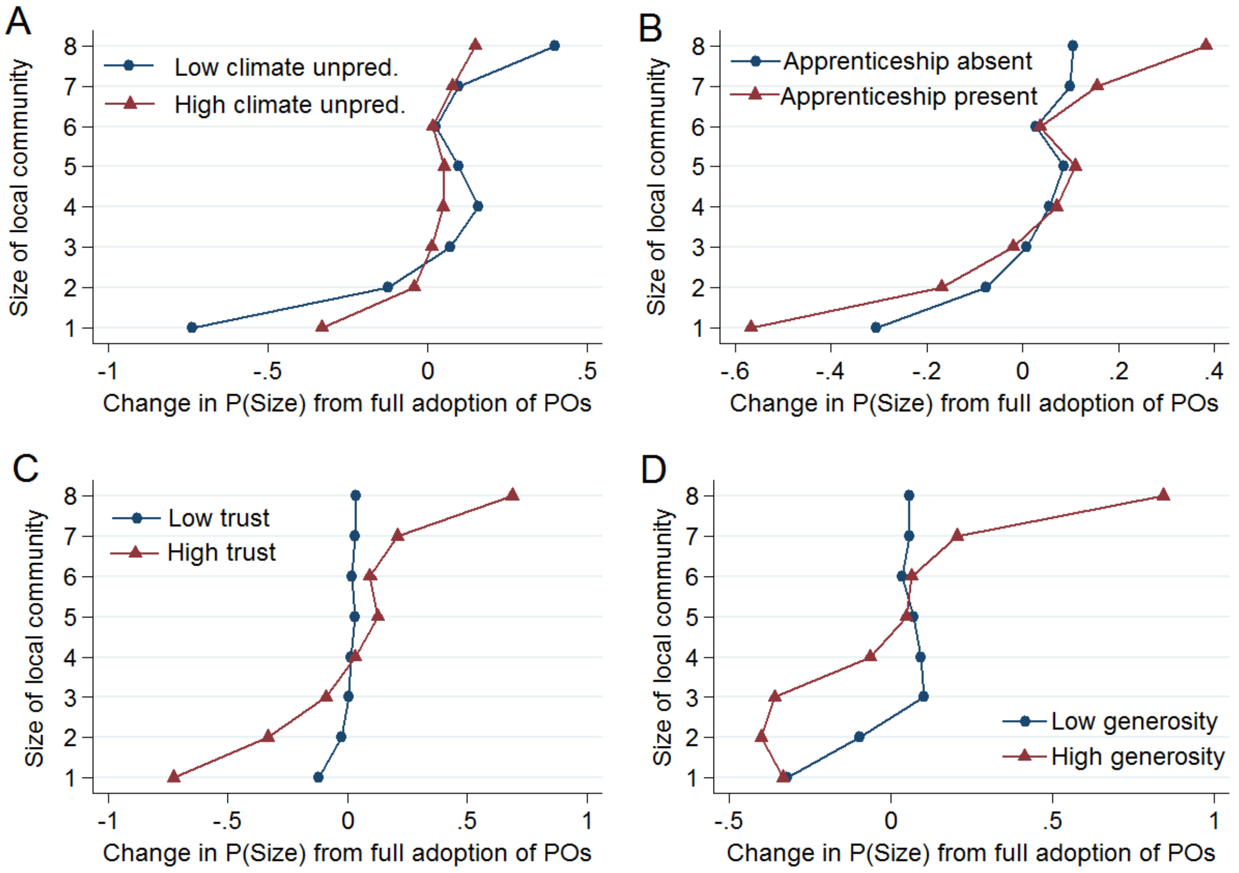
Supplementary Figure 10. Societies included in table 1. Map data: Google Maps

Supplementary Figure 11. Coefficients in quantile regression.
Notes: Coefficients in quantile regression. Dashed lines represent OLS coefficient, Continuous lines represent the coefficients from the quantile regression. In the regression we do not include the region controls because, if included, the regression would not converge.)



Supplementary Figure 12. Heterogeneity of impact of POs.

Notes: In this figure, we analyze the predictions from the comparative statics of our model. We use the estimation of table 5. (A) We plot the average of the marginal effects of "% within PO", that is, the change in the probability of each one of the eight size categories as this variable goes from 0 to 1. To show how this impact varies with climate unpredictability, we set this variable to the minimum and maximum values in the sample used in the estimation. (B / C / D) Analogous to graph A.



Supplementary Table 1. Descriptive statistics of continuous variables in table 1.

Variable	N	Mean	St. Dev.	Min	Max
Size of local population	173	3.38	2.33	1.00	8.00
% presence	173	0.67	0.18	0.27	1.00
% within PO	173	0.12	0.20	0.00	1.00
% presence x % within PO	173	0.10	0.16	0.00	0.64
Amphibian richness	173	15.64	15.87	1.00	102.12
Bird richness	173	202.72	109.90	2.75	528.85
Mammalian richness	173	73.12	41.77	2.75	196.00
Vascular plant richness	173	1817.97	879.41	2.75	4212.26
Distance to coast (km)	173	445.50	406.45	0.36	1697.74
Elevation	173	849.62	820.42	0.00	4345.04
Slope of terrain	173	2.73	2.90	0.00	14.13
Absolute latitude	173	31.71	18.95	0.00	71.00
Average temperature	173	13.18	9.81	-12.32	27.64
Year of ethnographic record	173	1880	212	-800	1965

Supplementary Table 2. Presence and intensity of agriculture

Category	**Number of societies**
Complete absence	66
Casual	8
Shifting cultivation	38
Semi-intensive, horticulture-like	5
Intensive with techniques (e.g., fertilization, crop rotation)	38
Intensive, dependent upon irrigation	18
Total	173

Supplementary Table 3. Type of Settlement

Category	**Number of societies**
Fully nomadic bands	19
Semi-nomadic (fixed winter settlement)	40
Semi-sedentary (shifting between fixed seasonal settlements)	27
Fixed settlements shifted every few years	3
Neighborhoods of dispersed family homesteads	18
Separate hamlets forming a community	6
Established Villages or towns	51
Complex settlements: Nuclear town with satellite homesteads/hamlets	9
Total	173

Supplementary Table 4. Regions

Region	**Number of societies**
Southwestern U.S.A.	24
Northwestern U.S.A.	13
Eastern Europe	12
Indian Subcontinent	9
West Tropical Africa	9
Western Canada	9
Western South America	7
East Tropical Africa	6
Northern South America	6
Papuasia	6
Brazil	5
Eastern Canada	5
Malasia	5
Southern Africa	5
Subarctic America	5
Australia	4
Russian Far East	4
Southeastern Europe	4
West-Central Tropical Africa	4
Siberia	3
South Tropical Africa	3
Southern South America	3
Western Asia	3
13 Other regions	19
Total	173

Supplementary Table 5. Year of ethnography

Year	**Number of societies**
800 BC	1
1520	1
1530	1
1750	2
1800 - 1819	2
1820 - 1839	1
1840 - 1859	18
1860 - 1879	28
1880 - 1899	23
1900 - 1919	24
1920 - 1939	31
1940 - 1959	36
1960 - 1965	5
Total	173

Supplementary Table 6. Selection on observables versus selection on unobservables

	1	2	3	4
	Local community size	Total population size	Population density	Cultural complexity
% presence	-0.143 (1.166)	-0.564 (1.997)	-0.669 (0.939)	-0.029 (5.155)
% presence x % within PO	5.238*** (1.592)	5.818*** (2.241)	3.159*** (1.202)	14.887*** (5.650)
All controls of table 1?	Yes	Yes	Yes	Yes
Observations	173	153	125	125
R-Square	0.771	0.819	0.409	0.888
Oster delta (R^2 max = 1)	0.41	0.52	0.71	0.93
Oster delta (R^2 max = .95)	0.51	0.69	0.98	1.69

Robust standard errors are used in all regressions and are displayed in parentheses. *** indicates p-value<0.01, ** indicates p-value<0.05, * indicates p-value<0.1 on a two-side t-test. The Oster delta is computed for the interaction term "%presence x within PO". This Oster delta assumes a linear model, so we estimate columns 1 and 3 using OLS (in table A8, we use the more appropriate ordered probit estimation for these dependent variables).

Supplementary Table 7. Instrumenting the presence of technology

Stage:	1st stage	2nd stage		2nd stage	
Dependent variable:	% presence	Local community size		Population size	
	1	2	3	4	5
% presence		4.793 (4.219)	2.958 (3.791)	12.982** (6.492)	6.221 (5.114)
% presence x % within PO			5.475*** (1.217)		10.865*** (2.669)
Kinship Tightness	0.061 (0.044)				
Sex differentiation	0.119** (0.053)				
Same controls as in table 1?	Yes	Yes	Yes	Yes	Yes
Region dummies and resource endowment controls?	No	No	No	No	No
Observations	194	194	194	160	160
Cragg Donald f-test first stage (p-value)		9.55*** (0.008)	9.62** (0.022)	8.884** (0.012)	8.085** (0.044)
Hansen test (p-value)		0.309 (0.578)	0.707 (0.702)	3.438* (0.064)	3.843 (0.146)

Robust standard errors are used in all regressions and are displayed in parentheses. *** indicates p-value<0.01, ** indicates p-value<0.05, * indicates p-value<0.1 on a two-side t-test. In columns 3 and 5, we use "kinship tightness x % within PO" and "sex differentiation x % within PO" as instruments for "% presence x % within PO". We drop the controls of geographical region because the local geographical variation in our instruments is not high. Given that we rely on inheritance from ancestral societies, the societies occupying a particular region tend to share the several cultural traits from their common ancestor. We drop the resource endowment variables in order to avoid data loss and to avoid small sample bias in the IV estimation (results are consistent if we include these controls).

Supplementary Table 8. Instrumenting the percentage within PO

Stage:	1st stage	2nd stage	2nd stage
Dependent variable:	% presence x % within PO	Local community size	Total population
	1	2	3
% presence	0.278*** (0.082)	-0.408 (1.223)	-1.316
% presence x % within PO		6.606* (3.751)	17.70*** (6.071)
% presence x Distance from Africa	-5.37e-06** (2.73e-06)		
Same controls as in table 1?	Yes	Yes	Yes
Region dummies?	No	No	No
Observations	173	173	153
Cragg Donald f-test first stage (p- value)		5.228** (0.022)	5.790** (0.016)

Robust standard errors are used in all regressions and are displayed in parentheses. *** indicates p-value<0.01, ** indicates p-value<0.05, * indicates p-value<0.1 on a two-side t-test. We use LIML in the estimations. We exclude region dummies because the variation of our instruments within regions is low.

Supplementary Table 9. Heterogeneity in the impact of POs

	Dependent variable: Size of local population					
	1	2	3	4	5	6
% presence	-3.24 (3.62)	-0.536 (1.590)	0.125 (1.241)	1.112 (1.694)	1.364 (1.701)	0.973 (2.679)
% presence x % within PO	12.28# (8.56)	2.629*# (1.547)	-0.811# (4.684)	0.066 # (2.920)	0.690 # (2.999)	1.898 # (3.231)
Climate unpredictability	-3.25 (4.06)					
Clim. unp. x % presence	4.75 (5.03)					
Clim. unp. x % pres. x % within PO	-10.66# (11.38)					
Apprenticeship		-0.553 (1.129)				
Appren. x % presence		0.237 (1.556)				
Appren. x % pres. x % within PO		3.336*# (1.860)				
Sex differentiation			not included			
Sex differ. x % presence			-0.078 (1.216)			
Sex differ. x % pres. x % within PO			6.955# (6.468)			
Trust				0.250 (0.179)		
Trust x % presence				-0.275 (0.314)		
Trust x % pres. x % within PO				0.768 # (0.644)		
Honesty					0.313 (0.271)	
Honesty x % presence					-0.622 (0.392)	
Hon. x % pres. x % within PO					1.148*# (0.652)	
Generosity						0.574** (0.289)
Gen. x % presence						-0.517 (0.426)
Gen. x % pres. x % within PO						0.598 # (0.615)
Resource endowment control?	Yes	No	Yes	No	No	No
All other controls of table 1?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	173	136	173	101	84	79
Pseudo R Square	0.354	0.335	0.353	0.239	0.281	0.299

Notes: Robust standard errors are used in all regressions and are displayed in parentheses. *** indicates p-value<0.01, ** indicates p-value<0.05, and *** indicates p-value<0.1 . # indicates p-value<0.01 for joint two tailed F-test of dsize/d%withinPO. To avoid losing excessive observations by using the SCCS variables, the sample of columns 2, 4, 5, and 6 allow at least 6 technologies with available information.

Supplementary Table 10. Robustness to other dependent variables

Dependent variable:	Total population			Population density			Cultural complexity		
	1	2	3	4	5	6	7	8	9
% presence	0.773 (2.253)	-0.564 (1.997)	-21.167*** (5.947)	0.789 (0.859)	-0.116 (0.877)	-6.213 (3.694)	4.660 (4.794)	0.313 (4.683)	-8.208 (13.050)
% pres. x % within PO		5.818*** (2.241)	24.996# (14.409)		3.584*** (1.101)	8.537# (5.423)		18.910*** (4.897)	64.518*** (18.426)
Climate unpredictability			-12.314** (5.659)			-1.256 (4.509)			-2.943 (14.188)
Climate unp. x % presence			29.596*** (8.307)			8.808 (5.519)			11.606 (20.434)
Climate unp. x % presence x % within PO			-27.183# (19.840)			-7.028# (6.976)			-61.207*** (23.719)
All controls of table 1?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	153	153	153	128	128	127	128	128	127
Pseudo R-Square (r-square)	0.801	0.819	0.835	(0.341)	(0.373)	(0.388)	0.807	0.828	0.835

Robust standard errors are used in all regressions and displayed in parentheses. *** indicates p-value<0.01 and # indicates p-value<0.01 for joint two tailed F-test of ∂size/∂%withinPO in columns 3, 6 and 9. "Population density" is the variable v1130 of the SCCS and is a categorical variable, with 1 equal to "less than 1 person per square mile", 2 equal to "1 - 4.9 persons per square mile", 3 equal to "5 - 24.9 persons per square mile", 4 equal to "25 - 99.9 persons per square mile", 5 equal to "100 - 499.9 persons per square mile", and 6 equal to "500 or more persons per square mile". High multicollinearity -- frequent in interaction models -- requires a joint test. "Cultural complexity" is the variable v158.1 of the SCCS, where they sum the scores of 10 variables that measure: writing and records, fixity of residence, agriculture, urbanization, land transport, money, density of population, political integration, social stratification, and specialization in metal working, weaving and pottery. In our case, we subtracted from the variable the last component of specialization. If we do not include density of population, results do not change. Columns 4, 5 and 6 use Ordered Probit, the rest OLS. For the columns 4 through 9, we use societies with information about at least 8 technologies in order to accommodate for the smaller sample size in these dependent variables (in table 6 below, we show that the results are robust to data stringency). Also, in columns 4 to 9 we do not use the set of dummies of region because approximately 20 observations a fully determined, and thus standard errors become suspect (if we include the region dummies the results do not change; if any, they become stronger).

Supplementary Table 11. Robustness to trade

	Dependent variable: Local community size					
	1	2	3	4	5	6
% presence	1.079 (1.235)	0.762 (1.267)	0.751 (1.243)	1.892 (1.421)	0.980 (1.227)	0.925 (1.259)
% presence x % within PO	4.237*** (1.530)	3.897*** (1.524)	3.978** (1.551)	2.630* (1.559)	4.261*** (1.532)	4.072*** (1.503)
"Intercommunity trade" dummies?	No	Yes	No	No	No	No
"Money" and "Credit" dummies?	No	No	No	Yes	No	No
"Importance of trade" control?	No	No	No	No	No	Yes
Resource endowment controls?	No	No	No	No	No	No
All other controls?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	130	130	124	124	131	131
Pseudo R-Square	0.367	0.379	0.368	0.422	0.367	0.375

We execute ordered probit regressions. Robust standard errors are used in all regressions and are displayed in parentheses. *** indicates p-value<0.01, ** indicates p-value<0.05, * indicates p-value<0.1 on a two-side t-test. To avoid data loss, in all regressions, we use societies with information about at least 8 technologies and drop resource endowments controls. We use four variables that proxy for trade: i) "intercommunity trade" (v1 of the SCCS) is a categorical variable that measures the extent to which intercommunity trade is a source of food (from "no trade" to "food imports present and contribute more than 50%"); ii) "presence of money" (v17) is a categorical variable with five categories (from "no media of exchange or money" to "indigenous coinage or paper currency"); iii) "presence of credit" (v18) is a categorical variable with four categories (from "Personal loans between friends or relatives" to "banks or comparable institutions"); iv) and "importance of trade" (v819b) is a continuous variable measures the percentage that trade contributes to subsistence in the society (i.e., food provision) (this variable is computed by the SCCS from using v1 and other five variables that provide categorical information on the extent that agriculture, fishing, gathering, animal husbandry and hunting contribute for subsistence; the mean is 8% with a maximum of 65%, a median of 5% and a 90th percentile equal to 25%). In column 2, we add "Intercommunity trade" as a control; in column 1, we use the same societies used in column 2. This allows us to cleanly assess the impact of the control variable on the impact of PO. The same is done in columns 4 and 3 for "money" and "credit" and in columns 6 and 5 for "importance of trade".

Supplementary Table 12. Robustness to political hierarchy

	1 Local community size	2 Population size	3 Population density	4 Cultural complexity
% presence	-0.190 (0.861)	-0.253 (2.053)	-0.669 (0.939)	-0.020 (3.350)
% presence x % within PO	3.790*** (1.428)	4.851** (2.490)	3.159*** (1.202)	8.964** (4.360)
Political hierarchy dummies?	Yes	Yes	Yes	Yes
Region dummies?	Yes	Yes	No	No
All other controls from table 1?	Yes	Yes	Yes	Yes
Observations	171	151	125	125
R-Square (Pseudo R-Square)	(0.358)	0.830	(0.409)	0.888

Notes: The variable "political hierarchy" is a categorical variable from the Ethnographic Atlas that indicates whether the society has political authority and if it does, the reach of this authority. The variable has five categories: 1 for "no political authority beyond the local community", 2 for "One level of hierarchy (e.g., petty chiefdoms)", 3 for "Two levels of hierarchy (e.g., large chiefdoms)", 4 for "Three levels of hierarchy (e.g., states), and 5 for "Four levels of hierarchy (e.g. large states)". In this table, we replicate column 2 of table 1 and columns 2, 5 and 8 of table 3, but with the addition of the control of political hierarchy. For "Population density" and "Cultural complexity", we do not use region dummies, and we use societies with information in at least 8 technologies in order to accommodate for the smaller sample sizes for these dependent variables. Robust standard errors are used in all regressions and displayed in parentheses. *** indicates p-value<0.01 on a two-side t-test. Columns 1 and 3 use OLS; columns 2 and 4 use Ordered Probit.

Supplementary Table 13. Robustness to available information on activities and regions

	Dependent variable: Size of local population (Ordered probit)					
	Number of activities				Societies per region	
	1	2	3	4	5	6
sample:	at least 7	at least 8	at least 9	at least 10	at least 2	at least 3
% presence	0.998** (0.472)	0.956* (0.525)	0.912 (0.590)	-0.047 (0.647)	0.013 (0.828)	-0.246 (0.832)
% presence x % within PO	3.230*** (0.715)	3.346*** (0.759)	3.201*** (0.837)	3.412*** (0.978)	4.185*** (1.243)	3.513*** (1.261)
All controls of table 1 included?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	430	393	330	269	166	154
Pseudo R-Square	0.287	0.286	0.295	0.306	0.342	0.3

We execute ordered probit regressions. Robust standard errors are used in all regressions and are displayed in parentheses. *** indicates p-value<0.01, ** indicates p-value<0.05, * indicates p-value<0.1 on a two-side t-test.

8 List of references cited in the supplementary information

Argote, L., and E. Miron-Spektor 2011 "Organizational Learning: From Experience to Knowledge." Organization Science 22 (5): 1123-1137

Arora, A., W. Cohen, and C. Cunningham 2018 "Inventive Capabilities in the Division of Innovative Labor" NBER working paper No. 25051.

Brahm, F., Tarziján, J. 2012. "The Impact of Complexity and Managerial Diseconomies on Hierarchical Governance", Journal of Economic Behavior and Organization, Vol 84, Nº 2, p. 586-599

Bowles, S., Choi, J. K., & Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. Journal of theoretical biology, 223(2), 135-147

Boyd, R. and P.J. Richerson 2005. The Origin and Evolution of Cultures. Oxford: Oxford University Press.

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

Boyd, R., Richerson, P. J., and J. Henrich, 2013. "The Cultural Evolution of Technology: Facts and Theories." In Cultural Evolution: Society, Technology, Language, and Religion, ed. Peter J. Richerson and Morten H. Christiansen.

Coase, R. 1937 "The Nature of the Firm." Economica 4 (16): 386-405.

Cohen, W. M., Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative science quarterly, 128-152

Colwell, R. 1974. "Predictability, Constancy, and Contingency of Periodic Phenomena" Ecology 55 (5): 1148-1153.

Cordes, C., Richerson, P. J., McElreath, R., & Strimling, P. (2008). 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

Depetris-Chauvin, E., and O. Ozak 2017. "The Origins and Long-Run Consequences of the Division of Labor" Working paper.

Enke, B. 2019. "Kinship Systems, Cooperation and the Evolution of Moral Systems." The Quarterly Journal of Economics, 134 (2): 953-1019

Epstein, S. R. (1998). Craft guilds, apprenticeship, and technological change in preindustrial Europe. The Journal of Economic History, 58(3), 684-713

De la Croix, D., Doepke, M., and J. Mokyr. 2018. "Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Pre-Industrial Economy." The Quarterly Journal of Economics, 133(1), pp.1-70.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. Nature, 425(6960), 785-791.

Fairlie, R W., Hoffmann, F., and P. Oreopoulos 2014 "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom." American Economic Review, 104 (8): 2567-2591

Fogarty, L., Strimling, P., and K.N. Laland 2011. "The Evolution of Teaching." Evolution 65 (10): 2760-2770

Haun, D., and H. Over 2013 "Like Me" In Cultural Evolution: Society, Technology, Language, and Religion, ed. Peter J. Richerson and Morten H. Christiansen.

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

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

Holmstrom, B., and P. Milgrom 1994. "The Firm as an Incentive System." The American Economic Review 84 (4): 972-991.

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