



Working Paper no. **173**

The evolution of workplace control
Leadership, obedience and organizational performance

Stefano Dughera
University of Paris-Nanterre, University of Torino

December, 2019

The evolution of workplace control

Leadership, obedience and organizational performance

Abstract: We study the interactions between supervisors and workers via evolutionary game-theory. We develop a simple model where workers select their effort level and decide whether to comply or defect. Supervisors, in turn, have two different functions: first, they act as coordinators and convert team-effort into output; and second, they act as monitors and induce workers to comply. In doing both, they may either behave as “motivators” through charisma and persuasion (authoritative style), or as “punishers” through authority and control (authoritarian style). In our framework, motivators encourage independence-seeking and reward compliers through inspiration and engagement, while punishers take all relevant decisions and reprehend defectors when these are caught shirking. The message is that authoritative leadership improves both productivity and worker’s well-being. For this to happen, supervisors must improve their charisma, while workers must develop their decisional skills. When either of these conditions is unmet, a variety of welfare-depressing situations may emerge.

KEYWORDS: leadership · organizational theory · personnel management · evolutionary games · mass-behavior

JEL CODES: C73 · M54

1. Introduction

Motivating obedience is a major problem in organizations. Given that effort is costly for the employee and valuable but difficult to monitor for the employer, the former has an incentive to disobey the latter, who is forced to find mechanisms to secure organizational performance. The shirking version of efficiency wage theory (Shapiro and Stiglitz, 1984; Bowles, 1985) postulates that a system of credible threats based on high wages and tight monitoring will induce workers to increase their labor effort. In doing so, it suggests that the use of authority is required to obtain what Nelson and Winter (1982: 110) call a «comprehensive truce in intraorganizational conflict». This is consistent with Coase's (1937: 387) description of the firm as a mini-command economy and even more with Marx's (182: 549). In his words:

The technical subordination of the worker to the uniform instruments of labour and the particular composition of the working group, consisting as it does of individuals of all sexes and ages, gives rise to a barrack-like discipline, which is elaborated into a complete system in the factory, bringing the previously mentioned labour of superintendence to its fullest development, thereby dividing the workers into manual labourers and overseers, into the private soldiers and the N.C.O.s of an industrial army.

It is our contention, however, that the specific mechanism whereby a supervisor secures her supervisees' obedience through authority and control is just one in a series of possible others. Swank and Visser (2007) and Kamphorst and Swank (2013, 2016), for instance, show that the allocation of tasks among workers of different abilities likely have motivational effects. In addition, when various psychological mechanisms are allowed into the picture, supervisors may affect obedience in several ways, from discriminating employees to boost her coworkers' self-confidence (Kamphorst and Swank, 2016), to ostracize workers and leverage on their reputational concerns (Kamphorst and Swank, 2013). Moreover, as those in supervisory positions have multiple functions in organizations (Mintzberg, 1971), emphasizing their role of monitors while neglecting that of, say, coordinators, may lead to disregard the interrelation

between these different roles. For instance, entrusting employees with decision-making power seems largely inconsistent with a disciplinary relationship based on authority and control. Since monitoring crowds out intrinsic motivation by signaling distrust (Frey, 1993; Chang and Lai, 1999), the idea is that it likely discourage workers to use their skills and take autonomous decisions.

The follower-leader literature in human resource management is rife with partially overlapping definitions of alternative leadership styles (Tannenbaum and Massarik, 1957; Bass, 1985; Yukl, 2010; Cunliffe, 2011). Despite their peculiarities, they are all characterized by a parsimonious use of authority and a greater emphasis on skills such as empathy, charisma and the likes. Management scholars have been urging leaders to change their supervisory attitude ever since Taylor's principles of work organization have been challenged by the emergence of alternative practices such as multitasking, horizontal communication, team-working and the likes (Boltanski and Chiappello, 2005). The normative character of these recommendations is well-captured by the following quotation:

Traditional bosses have little to do in a reengineered environment. Managers have to switch from supervisory roles to acting as facilitators, as enablers and as people whose jobs are the development of people and their skills so that those people will be able to perform value-adding processes themselves (Hammer and Champy, 1993: 82).

In a similar fashion, attachment theorists in the field of educational psychology have identified different parenting styles, the most effective of which have been conceptualized as authoritative and authoritarian (Baumrind, 1971; Darling and Steinberg, 1993). A parenting style is a theoretical construct describing the way in which genitors manage their child rearing activity, including the setting of rules and routines and the practices to motivate obedience with the latter. In this framework, authoritative parents encourage independence-seeking, prefer using non-punitive measures and are highly responsive to her children's needs and requests. Authoritarian parents, on the other hand, tend to overcontrol their pupils, punish disobedience

and provide little motivation for the rules they impose. Key to this distinction is the disciplinary differences between punishing those who disobey (the “stick” incentive) or encouraging those who obey (the “carrot” incentive). This clear-cut characterization will be handy for the analysis to follow.

The purpose of this study is to incorporate the insights from these neighboring literatures into a simple economic framework. The idea is to investigate the effect of different leadership styles of productivity and workers’ well-being, as to develop an economic theory of supervision beyond the authority-oriented approach of efficiency wage theory. The core message is that it is the *style*, not the *quantity* of supervision which determines morale and organizational performance, as those in supervisory positions may be either perceived as role-models to look up to or as tyrants to fight back. To provide with a simple formalization of this idea, we develop an evolutionary version of the classical shirking-model (Shapiro and Stiglitz, 1984; Bowles; 1985) where boundedly rational workers and supervisors interact in production. The former select their effort level and decide whether to cooperate or to defect, while the latter choose between two alternative strategies: either they elicit workplace discipline by motivating compliers through inspiration and engagement (authoritative style), or they tighten surveillance to punish defectors when these are caught shirking (authoritarian style). In our framework, the key variable in the leaders’ strategic behavior is personal magnetism, as poorly charismatic supervisors have no choice but imposing their authority to secure obedience. The study we feel closest to ours is that of Chang and Lai (1999), which explicitly account for the drawbacks of authoritarian control and analyze the interaction between positive and negative incentives. Even in this refined version of the classical shirking mode, however, positive inducements are always reduced to the wage payment and no light is shed upon the coaching skills of supervisors as a key organizational variable. As in the remainder of this literature, «the multifaceted relation between superior and subordinate is [...] reduced to a contractual one»

(Swank and Visser, 2007: 731). We see our study as an attempt to overcome such limitation.

Moreover, we make two additional contributions to the literature on intraorganizational conflict. First, we account for the multiple functions that supervisors play in organizations and allow the choice of the leadership style to affect the way in which leaders act as monitors and coordinators alike. In line with the insights from parenting theory (Baumrind, 1971; Darling and Steinberg, 1993), we assume that authoritative supervisors promote independence-seeking and allow their workers to take autonomous decisions, while authoritarian supervisors compel their employees to follow step-by-step procedures. The choice of the leadership style, in this respect, is driven by the workers' decisional skills: when employees take better decisions than their supervisors, authoritative leadership enhances productivity (with positive repercussions on workers' well-being); when employees take worse decisions than their supervisors, authoritarian leadership enhances productivity (with negative repercussions on workers' well-being). We intend this contribution as a preliminary exploration of the issue and leave further refinements for future research.

Second, we include a team-element in the model and assume that each supervisor oversees a team of two workers. This simplifying assumption allows us to inquire into some of the implications of mixed morale among co-workers and identify a novel effect in the literature on worker effort. The intuition is simple: by deciding to shirk, defectors impose a negative externality upon compliers, which takes the form of a psychological distress arising from the sense of procedural injustice. We dub this mechanism the "justice effect". Most interesting for our purposes is the possibility that a complier may start shirking to alleviate this psychological disutility. In this framework, high-performing monitoring technologies may impede this "bad company" effect to escalate. By successfully helping to punish defectors, in fact, they may increase the sense of procedural justice in organizations (Alge, 2001) and thus prevent compliers to start defecting. This novel insight adds on previous ideas submitted in the

literature which investigate the relationship between monitoring and effort (Frey, 1993; Chang and Lai, 1999).

The motivation of the study is straightforward. While personnel psychologists and management scholars have long insisted on the importance of getting the right fit between individual, job and managerial characteristics (Kristof-Brown, Zimmerman and Johnson, 2005), recent evidence on the correlation between management practices, productivity and workers' well-being suggests the need of further theoretical investigations in the field. Lazear, Shaw and Stanton (2015), for instance, estimate an up to 50% increase in the productivity of a worker moving from a low-quality to a high-quality boss, while the study run by Google in 2009 showed that high-productivity managers had subordinate engineers who scored highly on performance and retention (Garvin, Wagonfeld and Kind, 2013).

Beyond this organization-specific purpose, we believe that our contribution may find applications in several other fields, from child-rearing, crime deterrence, international relations and so on. As a rough example, imagine a situation where a population of police and social services officers deal with one of potential offenders. To discourage disobedience, the State may either invest in the police department (the "stick" incentive) or in other social services, such as adult education, job training, subsidized housing and the likes (the "carrot" incentive). From the citizens' viewpoint, the choice to avoid deviant behaviors is driven by two different effects: by the likelihood of being caught (increasing in the level of social control) and by the (potentially ineffective) consultancy that the State provides to encourage prosociality. In a standard shirking framework, this latter element would be normally reduced to the monetary incentives in formal labor markets. Conversely, our study highlights the importance of inspiration and guidance as key supports to the behavioral development of individuals.

We find that motivating obedience through charisma and persuasion has positive effect on both productivity and worker's well-being. The key requirement for this situation to occur

is twofold. On the one hand, supervisors must develop their social and relational skills, as poorly charismatic leaders are forced into authoritarian behaviors to counterbalance their inability as motivators. On the other hand, workers must develop their decisional skills, as authoritative supervisors are assumed to encourage independence-seeking and autonomous decision-making. The motivating program imposed on workers, in fact, may be distorted by the latter inability to take autonomous decisions and a command-and-control framework restored to avoid miscoordination. Continuing education and on-the-job training for both workers and supervisors may have a role in avoiding this two welfare-depressing situations to occur.

The remainder of the paper is organized as follows. Section 2 presents the model's main assumptions, section 3 the replicator dynamics. In section 4, we analyze the multiple configurations emerging from different sets of parameters' values. In section 5, we discuss the welfare properties of the model and comment on the possible emergence of Pareto-inferior situations which depress both productivity and workers' well-being. Section 6.

2. The model

2.1 Assumptions

Consider a model economy where large populations of supervisors and workers interact in production. At each moment in continuous time there are many random encounters between members of the two population. To give a team-flavor to the model, we assume that each supervisor oversees a team of two workers¹. By doing so, we provide a simple framework to analyze some of the implications of mixed morale among co-workers. In each encounter,

¹ Of course, this is a simple abstraction to capture the collective nature of production and analyse some of the team-problems that it entails. In an evolutionary game-theoretic framework like ours, the main implication of considering larger teams relates to the frequency with which players experience a given payoff, and thus, to the numerosity and topological properties of the stationary points of the dynamics. The main messages of the of the model, however, carry over for any of these quantitative variations. For a model displaying a similar structure of interactions, see Antoci, Fiori Maccioni and Sacco. (2017).

workers exchange labor services for a uniform wage $w^H > 0$. As in Shapiro and Stiglitz (1984), we restrict our attention to the case in which they have only two effort-choices: either they “comply” and provide labor effort at a private cost $E > 0$ (strategy C) or they “defect” and exert no effort (strategy D). Since working is costly and the benefit to wage income applies to all workers, employees have an incentive to shirk.

Supervisors, in turn, have two different functions: first, they act as coordinators and convert team-effort into output; and second, they act as monitors and compel workers to comply. In doing both, they may either behave as “motivators” through charisma and persuasion (strategy M) or as “punishers” through authority and control (strategy P). In particular, we assume that punishers command their workers to follow step-by-step procedures and impose monitoring to detect counterproductive work behaviors. In doing so, they discourage independence-seeking and crowds out intrinsic motivation by signaling distrust (Frey, 1993; Chang and Lai, 1999). This is the kind of supervisory attitude which is usually modelled in the shirking version of efficiency wage theory (Shapiro and Stiglitz, 1984; Bowles, 1985) and corresponds to the authoritarian style in both leadership (Tannenbaum and Massarik, 1957; Bass, 1985; Yukl, 2010) and parenting (Darling and Steinberg, 1993) theory. Motivators, on the other hand, ask their employees to develop their skills and adapt to local information. To avoid breaching the implicit psychological contract with their team, we assume that they use non-punitive measures to obtain their workers’ cooperation. More specifically, they try stimulating the compliers’ intrinsic motivation through charisma and psychological support. This corresponds to the authoritative style in parenting theory (Darling and Steinberg, 1993) and, less sharply, to a variety of non-authoritarian styles in leadership theory (Tannenbaum and Massarik, 1957; Bass, 1985; Yukl, 2010; Cunliffe, 2011).

Following Shapiro and Stiglitz (1984), we assume that workers contribute one unit of labor if they do not shirk and zero otherwise. Hence, we denote by $R^P > 0$ the productivity of

a complier meeting a punisher, and by $R^C \geq 0$ the productivity of a of a complier meeting a motivator, where $R^i, i = P, C$ are quality parameters that convert effort into output. The superscripts P (punisher) and C (complier) indicates who manages the decision-making process. When the supervisor is a punisher, she makes all relevant decisions. Conversely, when the supervisor is a motivator, workers are given the right to take initiatives. As the supervisor knows her ability as decision maker but cannot assess but imperfectly her workers', she is ex ante unable to foresee if delegating decision-making will improve or rather hinder organizational performance. We introduce the following convention: if $R^C > R^P$, we say that workers are high-skilled (they take better decisions than their supervisors); if $R^C < R^P$, that workers are low-skilled (they take worse decisions than their supervisors). Needless to say, as defectors exert no effort, their productivity is null regardless of the kind of supervisor they are paired with.

Allowing for situations where workers “take better decisions” than their likely more experienced supervisors may sound counterintuitive. On second thoughts however, the reasons for this to happen are multiple. First, the workload of business leaders is often overwhelming, up to the point where they get to concede but cursory attention to each of their multiple tasks. Workers, on the other hand, are normally in charge of fewer duties and should be able to pay greater attention to their daily decisions and routines. In his study on managerial work, Mintzberg (1971) highlights how frenetic the typical day of a manager can be, thus providing support to the idea that delegating decision-making may improve organizational performance². Second, business leaders may have “the vision” and the skills to coordinate it throughout, but may lack the idiosyncratic knowledge to decide on task-specific matters. Third,

² To provide with a simple formalization of this intuition, consider the following. Assume that the above productivities now write $R^P = S^P(1 - \tau)$ and $R^C = S^C$, where $0 < \tau < 1$ indicates the amount of time that supervisors spend in decision-making, while $S^P > 0$ and $0 < S^C < S^P$ measure the supervisor's and the worker's decisional skills respectively. When $\tau > (S^P - S^C)/S^P$, supervisors are overloaded and delegating decision-making improves organizational performance, regardless of the fact that supervisors are better-skilled than employees.

as workers are closer to production, autonomous decision-making reduces costly communication.

To provide with a broader intuition for this modelling strategy, consider a parent helping her kid choosing the faculty to enroll in. The parent's payoff is a positive function of her child's academic achievement, which may be loosely defined as the kid's "productivity" in forming human capital. The genitor believes that choosing a given faculty (say, law) is the best option for her kid. Of course, this belief is likely biased by the parent's education and occupational history, as she likely knows the outcome of pursuing a given path but ignores the others. For instance, she may have studied law and acquired knowledge of the market for legal professions, but remained oblivious of the functioning of other labor markets. The child, on the other hand, would prefer going for another option (say, philosophy). In this case, the parent may either compel her kid to "follow the order she is given because she said so" or may entrust the latter with the autonomy to pursue her own decision. If the kid is committed to her choice, this may eventually yield higher payoffs to the genitor and the kid alike ($R^C > R^P$). Conversely, if the kid takes bad decision or follow them poorly, forcing the latter to pursue the genitor's preferred path may be beneficial, if not for both, at least for the parent ($R^C < R^P$).

To entice commitment from supervisors, we assume that organizations allow the supervisory compensation to depend on team-performance³. As the share of revenue allocated to supervisors is assumed to be equal across the two regimes, we can proceed as if supervisors paid wages and got the residual reward. In this framework, the supervisory payoff turns out to be a measure of firm's profitability. To ensure the existence of positive profits, we assume that $R^P > w^H$. In words, when supervisors decide to handle the decision-making process, the

³ More realistically, the supervisory compensation may consist in a uniform fixed wage plus a premium which varies with team performance. From a game-theoretic viewpoint, however, we are interested in analysing the payoff difference of a strategy with respect to its alternative strategy. Hence, we can focus on the variable premium only.

organization is profitable.

As anticipated, workplace discipline is maintained differently in the two supervisory modes. While motivators use their charisma to incentivize compliers, punishers reprehend defectors by monitoring team performance. To model the supervisors' motivational skills, we assume that the psychological benefit of providing effort under the lead of a motivator is given by $qB \geq 0$, where $B > 0$ measures the workers' intrinsic motivation and $0 \leq q \leq 1$ indexes the level of the supervisor's charisma. The intuition is that workers need good bosses to stimulate their commitment and that this stimulus is all the more effective the more charismatic is the supervisor. As loafing workers disregard work morale, they only benefit from the expected wage payment and are thus insensitive to how charismatic their supervisor may be.

As concerns punishers, we follow Chang and Lai (2002) and assume that when a defector is caught shirking by a punisher, her wage falls down to w^L , where naturally, $0 \leq w^L < w^H$. Hence, $w^H - w^L \equiv W > 0$ can be referred to as the premium for high-effort. In addition, we assume that the punishers' monitoring capability varies with the composition of her team. When both workers defect, team production is null, monitoring is perfect, and the entire team receives the punishment w^L . Conversely, when one worker plays C and the other plays D , the supervisor cannot discern but imperfectly who is shirking in her two-members team. In this case, monitoring is imperfect, the hard worker receives the premium w^H and the shirkers' expected revenue is given by $(1 - p)w^H + pw^L$, where $0 \leq p \leq 1$ measures the exogenously given efficiency of the monitoring technology. Here, $-pW$ represents what Frey (1993) calls the "disciplining effect", which postulates that more intense monitoring and higher wages correlate positively with workers cooperation. In our framework, this effect varies negatively with the density of the low-effort group, as it is strong when both workers defect and weak when one worker defects and the other complies.

Moreover, we assume that unilateral shirking inflicts a negative externality upon

compliers, as coworking with a defector generates a sense of procedural injustice. In this view, high-performing monitoring technologies may reduce the compliers' psychological distress. We dub this positive effect the "justice effect". Support to these hypotheses comes from the field of industrial psychology. In particular, Alge (2001) analyzes the impact of computer surveillance on employees' morale and find that increasing control in organizations improves the sense of procedural justice among co-workers. In formal terms, we allow $K(p) \geq 0$ to measure the intensity of the negative externality, and we further assume that $K'(p) < 0$. To keep things simple, we allow $K(p)$ to take the tractable form $K(p) = K(1 - p)$. In words, when monitoring is perfect, the distress from coworking with unpunished defectors disappears. Hence, since motivators adopt no monitoring, a complier meeting a motivator and a defector incurs in the full psychological cost K , while this cost falls down to $K(1 - p)$ when she is matched with a punisher and a defector.

Finally, we incorporate the idea expressed by Frey (1993) and then modelled by Chang and Lai (1999) whereby monitoring depresses work morale, as it crowds out intrinsic motivation by signaling distrust. To keep things simple, we assume that compliers suffer a psychological distress $H(p) \geq 0$ when meeting a punisher, regardless of the composition of their team. In line with Chang and Lai (ivi), we further assume $H'(p) > 0$, as the crowding out of intrinsic motivation increases when monitoring becomes more intense. As before, we assume that $H(p)$ takes the tractable form $H(p) = Hp$.

To sum up, our model postulates that monitoring have three different effects on workers cooperation. The first is positive and corresponds to the standard disciplining effect predicted by the shirking version of efficiency wage theory (Shapiro and Stiglitz, 1984; Bowles, 1985). The second is newer in the literature on workers' effort and claims that employees will cooperate more when monitors successfully punish defectors, as this improves the sense of procedural justice within organizations (Alge, 2001). The third is negative and corresponds to

the crowding out effect theorized by Frey (1993). The combination of these mechanisms is ambiguous and depends on whether the crowding out effect dominates the disciplining and the justice effect or vice-versa. In the former case, as put forwards by Chang and Lai (1999), the wage and the monitoring incentives can be treated as complements; in the latter, as substitutes.

2.2 Payoffs

From the above discussion, it should be clear that the workers' payoffs depend on how the supervisors' choice of the leadership style interacts with the coworkers' decision to comply or defect—for a summary of the parameter's interpretation, see table 1. Hence, the workers' "incentives to comply" may vary across each of the four combinations between team composition and supervisory modes. When the payoff-difference between complying and defecting is positive in one of these combinations, we say that the incentives to comply in that combination are positive as well; when they are negative, that the incentives to comply are negative as well. The workers' payoff matrix is given by:

	<i>CM</i>	<i>DM</i>	<i>CP</i>	<i>DP</i>	
<i>C</i>	$w^H - E + qB$	$w^H - E + qB - K$	$w^H - E - Hp$	$w^H - E - Hp - K(1 - p)$	(1)
<i>D</i>	w^H	w^H	$(1 - p)w^H + pw^L$	w^L	

The choice of the supervisory regime, in turn, is driven by two elements: by the supervisors' capability of eliciting work effort, and by the workers' decisional skills—for a summary of the parameter's interpretation, see table 1. The supervisors' payoff matrix is given by:

	CC	$CD = DC$	DD
M	$2(R^C - w^H)$	$R^C - 2w^H$	$-2w^H$
P	$2(R^P - w^H)$	$R^P - 2w^H + pW$	$-2w^L$

(2)

Recalling that the supervisors' payoffs are a proxy for firm profitability in this model, we see that organizations make operating losses when their leaders fail to motivate obedience, regardless of the latter choice of the leadership style. From the third column of matrix (2), in fact, we see that team-production is null when both workers play D . However, since motivators avoid punishing defectors by lowering their wage to w^L , their losses are greater than the punishers'. In this case, punishing is always more rewarding than motivating. A key implication of our model, thus, is that supervisors may myopically choose to play P to self-protect from the risk of counterproductive work behaviors. If defecting can be supported as an equilibrium strategy, this may create a welfare-depressing situation where shirking and punishing reinforce each other.

Parameter	Interpretation
$w^H > 0$	High wage
$0 < w^L < w^H$	Low wage
$E > 0$	Effort cost
$0 \leq q \leq 1$	Motivators' charisma
$B > 0$	Worker's intrinsic motivation
$K > 0$	Sense of procedural injustice
$H > 0$	Crowding out of intrinsic motivation
$0 \leq p \leq 1$	Monitoring efficiency
$R^P > 0$	Supervisors' decisional skills
$0 < R^P \stackrel{>}{<} R^C$	Workers' decisional skills

Table 1. Interpretations of the payoffs

At each moment, we denote by $0 \leq x \leq 1$ the share of workers who choose to comply and by $(1 - x)$ the share of workers who choose to defect. Similarly, we denote by $0 \leq y \leq 1$ the share of supervisors who choose to motivate and by $(1 - y)$ the share of supervisors who choose to punish. Recall that each supervisor leads a team of two workers. Hence, the probability that a complier is matched with a complier and a motivator is xy , with a defector and motivator is $(1 - x)y$, with a complier and a punisher is $x(1 - y)$ and with a defector and a punisher is $(1 - x)(1 - y)$. Accordingly, the probability that a supervisor is matched with two compliers is x^2 , with one complier and one defector is $2x(1 - x)$ and with two defectors is $(1 - x)^2$. We calculate the expected payoff from complying and defecting from matrix (1), which are given, respectively, by:

$$\Pi^C = K(1 - p)x + (Hp + qB - Kp)y + Kpxy + w^H - E - Hp - K(1 - p)$$

$$\Pi^D = W(1 - p)x + Wy - (1 - p)Wxy + w^L$$

So that the payoff-difference between the two strategies writes⁴:

$$\begin{aligned} \Pi^C - \Pi^D &= (K - W)(1 - p)x + (Hp + qB - W - Kp)y + [Kp + (1 - p)W]xy + W \\ &\quad - E - Hp - K(1 - p) \end{aligned}$$

From which it is easy to derive the curve along which $\Pi^C - \Pi^D = 0$, whose equation writes:

$$y = \frac{E + Hp + K(1 - p) - W - (K - W)(1 - p)x}{Hp + qB - W - Kp + [Kp + (1 - p)W]x} \quad (3)$$

⁴ Inspecting further $\Pi^C - \Pi^D$, we see that improvements in the monitoring efficiency have an ambiguous effect on workers cooperation. This is given by the fact that $\partial(\Pi^C - \Pi^D)/\partial p \gtrless 0$. In particular, the effect is positive if $W(x - xy) + K(1 - y - x + yx) > H(1 - y)$, that is, if the disciplining and the justice effects outweigh the crowding out effect, and it is negative otherwise.

From matrix (2), we calculate the expected payoff from motivating and punishing, which are given, respectively, by:

$$\Pi^M = 2(R^C x - w^H)$$

$$\Pi^P = 2\{R^P x - w^L - W[x + (1 - p)(x - x^2)]\}$$

So that the payoff-difference between the two strategies writes:

$$\Pi^M - \Pi^P = 2\{(R^C - R^P)x - W[1 - x - (1 - p)(x - x^2)]\}$$

From which it is easy to derive the curve along which $\Pi^M - \Pi^P = 0$, whose equation writes⁵:

$$x = \frac{R^C - R^P + W(2 - p) - \{[R^C - R^P + W(2 - p)]^2 - 4W^2(1 - p)\}^{1/2}}{2W(1 - p)} \quad (4)$$

The next step is to describe how the system may evolve under alternative parametrizations.

3 Dynamics, equilibria and stability

We model the diffusion of the C and M -strategies in their respective populations via the standard replicator-dynamics derived by Taylor and Jonker (1978). The replicator-dynamics is a learning-by-imitation model which postulates that players are boundedly rational, they learn from each other, and they tend to adopt the strategy that performs better than the other, so that relatively successful behaviors are replicated, while unsuccessful behaviors are abandoned. In our framework, the idea that work morale and leadership inclinations are private information

⁵ Beware that there two roots of equation (4). Yet, one of them is always > 1 , as the opposite would require $p > 1$.

between workers and supervisors and cannot be signaled nor inferred when entering the employment relationship. The system's dynamics are given by:

$$\begin{cases} \dot{x} = x(1-x)(\Pi^C - \Pi^D) \\ \dot{y} = y(1-y)(\Pi^M - \Pi^P) \end{cases} \quad (5)$$

where \dot{x} and \dot{y} are the time derivatives of x and y respectively. Dynamics (5) is defined in the unit square $Q = [0, 1]^2$. As usual with replicator dynamics, all edges of the square are invariant⁶ and the four vertices $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$ where both populations are homogenous—they are both composed of one type only—are always stationary states. In addition, dynamics (5) may admit three other stationary points, the first and second of which correspond to the intersection, when existing, of the nullcline defined by (3) and the edges of Q where $y = 0$ and $y = 1$. In the first of these critical points—indicated as $(\hat{x}, 0)$, with $0 < \hat{x} < 1$ —supervisors play strategy M , \hat{x} workers play C and the remaining $1 - \hat{x}$ workers play D . In the second of these critical points—indicated as $(\tilde{x}, 1)$, with $0 < \tilde{x} < 1$ —supervisors play strategy P , \tilde{x} workers play C and the remaining $1 - \tilde{x}$ workers play D ⁷. The last stationary state admitted for by dynamics (5)—indicated as (\bar{x}, \bar{y}) , with $0 < \bar{x} < 1$ and $0 < \bar{y} < 1$ —is located in the interior of Q and corresponds to the intersection, when existing, of the nullclines defined by (3) and (4). In such state, all four types of players coexist. Observe that $\dot{x} = 0$ holds along the curve defined by (3) and along the edges of Q where $x = 0$ and $x = 1$, while \dot{y} holds along the curve defined by (4) and along the edges where $y = 0$ and $y = 1$.

⁶ Meaning that all trajectories starting from an initial pair $(x_0, y_0) = (1, \hat{y})$, $(x_0, y_0) = (0, \hat{y})$, $(x_0, y_0) = (\hat{x}, 0)$ and $(x_0, y_0) = (\hat{x}, 1)$ will lie on the side with $x = 1$, $x = 0$, $y = 0$ and $y = 1$ respectively, where $0 \leq \hat{x} \leq 1$ and $0 \leq \hat{y} \leq 1$.

⁷ All proofs are given in the Appendix, which is available from the author upon request.

To analyze the system's topological properties, we calculate the Jacobian matrix, which is given by:

$$J = \begin{bmatrix} (1 - 2x)(\Pi^C - \Pi^D) + x(1 - x)\alpha & x(1 - x)\beta \\ y(1 - y)\gamma & (1 - 2y)(\Pi^M - \Pi^P) \end{bmatrix}$$

Where we have defined: $\alpha \equiv [(K - W)(1 - p) + Kpy + (1 - p)Wy]$, $\beta \equiv Hp + qB - W - Kp + Kpy + (1 - p)Wy$ and $\gamma = -W(2 - 2p)x + R^C - R^P + W(2 - p)$. Evaluating J at each stationary point, we derive the dynamic's topological properties, which are summarized in the following Proposition:

PROPOSITION 1—*The stationary points $(0, 1)$, $(\tilde{x}, 1)$ and (\bar{x}, \bar{y}) , when existing, are either saddles or sources. In addition, the stationary point (\bar{x}, \bar{y}) exists if and only if no point is globally attractive. Finally:*

- (i) *The $(1, 0)$ equilibrium is attractive if and only if $p > \frac{E}{W-H}$ and $R^C < R^P$.*
- (ii) *The $(0, 0)$ equilibrium is attractive if and only if $p < \frac{K+E-W}{K-H}$.*
- (iii) *The $(\hat{x}, 0)$ equilibrium is attractive if $W > K$ and $R^P < R^C < R^P + \frac{1-\hat{x}-(1-p)(\hat{x}-\hat{x}^2)}{\hat{x}} \equiv \hat{R}$.*
- (iv) *The $(1, 1)$ equilibrium is attractive if $q > \frac{E}{B}$ and $R^C > R^P$.*

4. Dynamic regimes

From Proposition 1, it is easy to check that at most two equilibria may simultaneously attract. When a single stationary point is globally attractive, we say that the corresponding dynamic regime is monostable—see fig. 1—when two stationary points are locally attractive, we say that the corresponding dynamic regime is bistable—see fig. 2, 3 and 4—when no

stationary point is neither globally nor locally attractive, we say that the corresponding dynamic regime is cyclical or oscillatory—see fig. 5.

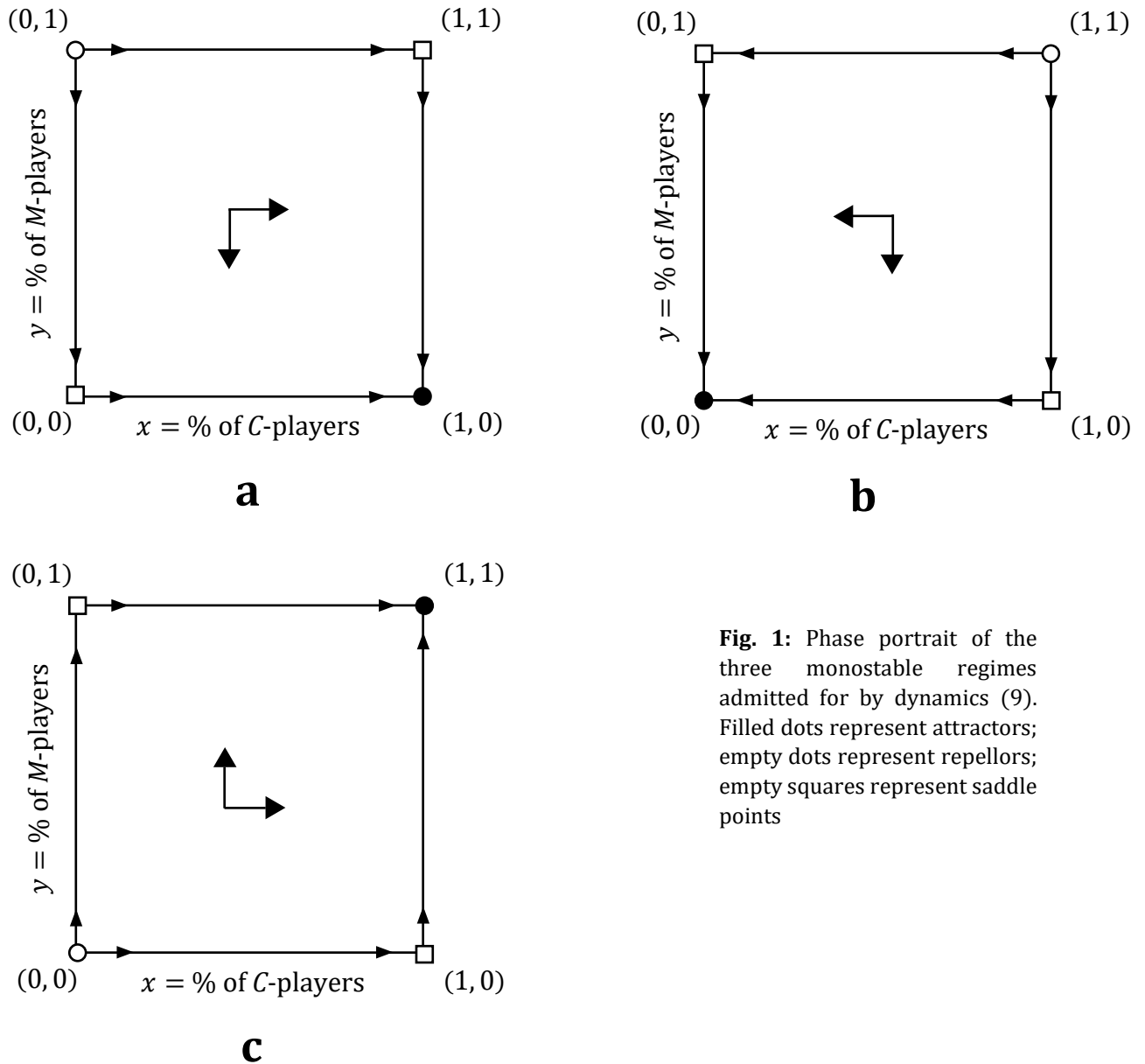


Fig. 1: Phase portrait of the three monostable regimes admitted for by dynamics (9). Filled dots represent attractors; empty dots represent repellers; empty squares represent saddle points

To refer to the potentially attractive equilibria in an intuitive way, we call the situation where all workers comply and all supervisors punish, “Discipline”, the scenario where all workers defect and all supervisor punish “Conflict”, the situation where all supervisors punish and workers randomize between complying and defecting, “Quasi-Discipline” and the scenario

where all workers comply and all supervisors motivate, “Cooperation”. Two remarks are worth drawing. First, the Cooperation equilibrium may only attract when workers are high-skilled, that is, when $R^C > R^P$. This is due to the admittedly imitating assumption whereby the choice of the leadership style affects the way in which supervisors act as monitors and coordinators alike. Under the assumption $R^P > R^C$, in fact, encouraging independence-seeking is counterproductive and the choice of the P -strategy is chiefly motivated by coordination reasons. Second, as both organizations and supervisors have negative payoffs in the Conflict equilibrium, this situation can be referred to as an organizational failure.

Inspecting further Proposition 1, we immediately see that dynamics (5) admits four monostable regimes, three bistable regimes and a cyclical or oscillatory regime. Hereafter, we analyze the most interesting of these configurations, i.e. when the system exhibits bistable or cyclical behavior.

4.1 The “Bad Company” regime

The first element which polarizes the system’s topology is the quality of the workers’ skills. When $R^C < R^P$, workers are low-skilled and are thus entrusted with no autonomy in decision-making. In addition, since the motivator’s losses are larger than the punishers’ when both lead a team of defectors, playing P is the supervisors’ dominant strategy in this regime. In this case, both the Cooperation and the Quasi-Discipline equilibria repel, while the Conflict and the Discipline equilibrium may either globally or locally attract. Hence, under the condition $R^C < R^P$, dynamics (5) admits a single bistable regime featuring the Discipline and the Conflict equilibria as local attractors—see fig. 2. The associated parametrization is given by:

$$R^C < R^P \text{ and } \frac{E}{W - H} < p < \frac{K + E - W}{K - H} \Leftrightarrow K > W \quad (6)$$

We dub this situation “Bad Company”, as the sense of procedural justice plays a major role

in this regime. Indeed, when compliance is diffused in the workers' population, the presence of sufficient C -players generates the critical mass to impede shirking to become endemic. Conversely, when the share of defectors is initially large, work morale will erode and shirking become widespread. The mechanism behind these diverging dynamics is well-captured by two different situations occurring at the team level.

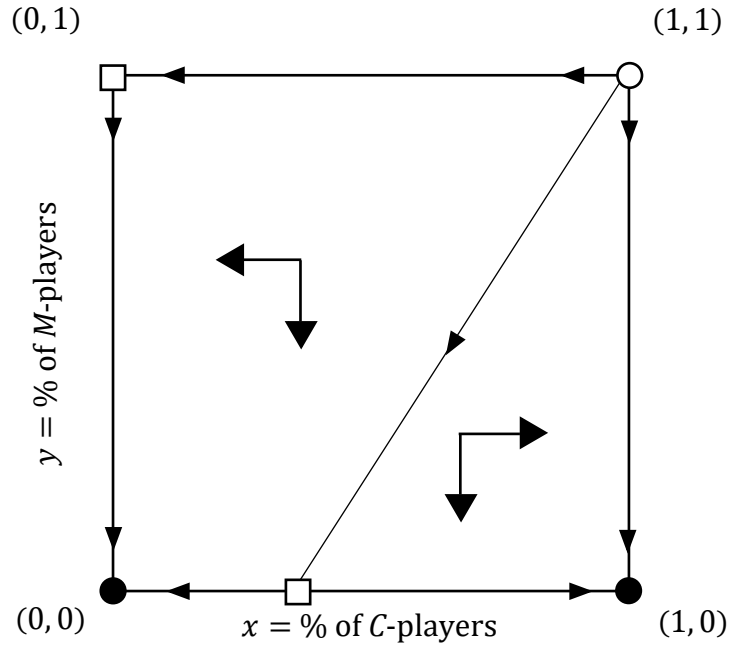


Fig. 2: Phase portraits of replicator dynamics (9) when condition (10) is satisfied (the “Bad Company” regime). Filled dots represent sinks; empty dots represent sources; empty squares represent saddle points. The trajectory connecting the $(1,1)$ source and the saddle located in the interior of $Q_{y=0}$ belong to the stable manifold of the saddle and it is the separatrix of the system.

When both workers comply, the detection probability is sufficiently high to discourage individual shirking. Formally, this is given by the fact that $p > E/(W - H)$. In this case, the incentives to comply are positive, as the expected benefit from complying $w^H - E - Hp$ outweighs the expected benefit from defecting $(1 - p)w^H + pw^L$. Conversely, when both team members defect, the incentives to comply are negative, as the sense of procedural injustice is too strong for either of the two workers to comply. This is given by the fact that $w^L > w^H - E - Hp - K(1 - p)$, so that each worker prefers receiving the low wage rather than feeling as the

“village idiot” who complies when her coworker shirks. In this case, the detection probability is too low to compensate compliers for their sense of procedural injustice. To overcome such situation, organizations may indifferently raise the premium for hard workers or increase the tightness of control. When condition (6) is satisfied, in fact, the wage and the monitoring incentives can be treated as substitutes, as monitoring improves procedural justice more than it crowds out intrinsic motivation. Formally, this can be seen from the fact that the compliers’ payoffs vary positively with the tightness of control, as $\partial[w^H - E - Hp - K(1 - p)]/\partial p > 0$ when $K > H$ ⁸.

4.2 The “Trading Places I” regime

When workers are high-skilled ($R^C > R^P$), they make better decisions than their supervisors, who have an incentive to delegate decision-making. In this case, the Discipline equilibrium is always repulsive, while the Conflict, the Quasi-Discipline and the Cooperation equilibrium may either globally or locally attract. Under the assumption $R^C > R^P$, the first element which polarizes the system’s behavior is the motivators’ charisma and, relatedly, their capability of motivating obedience. When $q > E/B$, we say that motivators are charismatic, when $q < E/B$, that they are not charismatic. Under the assumption $q > E/B$, in fact, the compliers’ payoffs are higher than the defectors’ when both are matched with a motivator and a complier, since $w^H - E + qB > w^H$.

When $R^C > R^P$ and $q > E/B$ are simultaneously satisfied, the model generates two different predictions, which we call “Trading Places I” and “Trading Places II” respectively. In

⁸ The fact that $K > H$ always in this regime can be proven as follows. First observe that a necessary condition for (8) to hold is that $(K + E - W)/(K - H) > 0$, which may either obtain if $E + H < W < E + K$ or if $E + K < W < E + H$. However, since we already know that $E/p + H < W$, $E + H < W < E + K$ always, which entails that $K > H$.

the first of these scenarios, the Cooperation and the Conflict equilibrium simultaneously attract—see fig. 3. The associated parametrization is given by:

$$R^C > R^P \text{ and } q > \frac{E}{B} \text{ and } p < \frac{K + E - W}{K - H} \quad (7)$$

As in the Bad company regime, the incentives to comply are negative for the individual complier who is matched with a punisher and a defector, since $w^L > w^H - E - pH - K(1 - p)$. Conversely, they are positive for the individual complier who is matched with a motivator and a fellow-complier ($q > E/B$). As in the Bad Company regime, initial conditions matter. When defectors are abundant in the workers' population, defecting becomes characterized by mass-behavior and it is sustained as an equilibrium strategy when both shirking and punishing become widespread. Conversely, when both compliers and motivators are initially abundant in their respective population, the system will eventually converge to the Cooperation equilibrium. This provides support to the idea anticipated earlier whereby intraorganizational conflict is self-reinforcing. Indeed, when peer-pressure is both negative and strong, authoritarian leaders are unable to break the vicious circle of misconducts, and yet they stick to their strategy as it pays off more than its alternative. The self-protective property of the P -strategy, in this case, provides supervisors with the wrong incentive, as continuing to play P yields the negative payoff $-2w^L$. As in the Bad Company regimes, raising the wage premium for hard workers promote organizational compliance. Depending on parameters' value, tightening control may also stimulate worker cooperation. In particular, this occurs if the negative externality of working with a loafing colleague is larger than the negative externality of being supervised by a distrustful boss, that is, if $K > H$.

Path dependency, here, is even stronger than in the Bad Company regime, as it is determined by the initial distributions of behavioral types across both populations. Incidentally, this is also the reason why we call this regime Trading Places I.

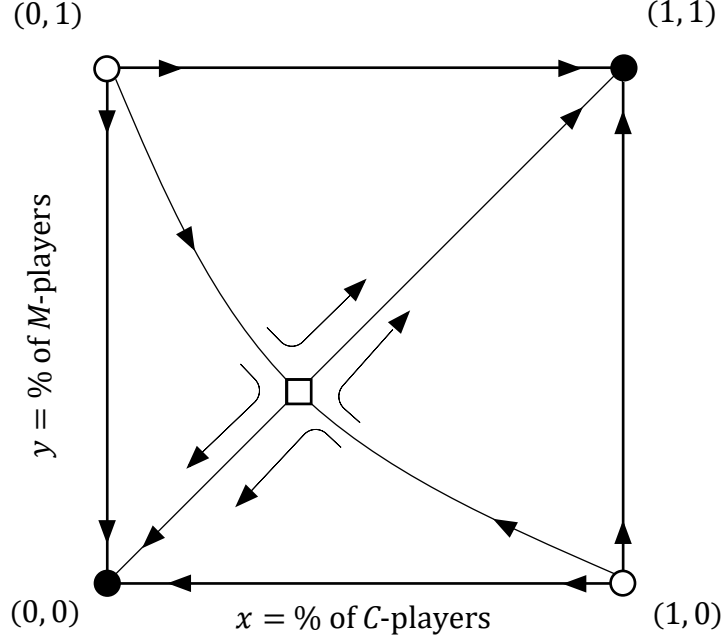


Fig. 3: Phase portraits of replicator dynamics (9) when condition (11) is satisfied (the “Trading Places I” regime). Filled dots represent attractors; empty dots represent repellers; empty squares represent saddle points and the two intersecting lines are the trajectories belonging to the stable and unstable branch of the internal saddle.

As in the homonymous movie, initial conditions matter, and individuals behave consistently with the kind of environment they are inserted in. When workers are punished, they retaliate and defect; when they are motivated, they reciprocate and comply. In the same vein, when norms are disregarded, supervisors decide to punish; when norms are obeyed, they decide to motivate.

4.3 The “Trading Places II” regime

When motivators are charismatic and workers are high-skilled, dynamics (5) admits a second bistable regime, where the Cooperation and the Quasi-Discipline equilibrium simultaneously attract—see fig. 4. We call this situation Trading Places II. The associated parametrization is given by:

$$0 < R^c < \hat{R} \text{ and } q > \frac{E}{B} \text{ and } W > K \Leftrightarrow \frac{K + E - W}{K - H} < p < \frac{E}{W - H} \quad (8)$$

The peculiarity of this regime is that workers are high-skilled, but “not too much” ($R^C < \hat{R}$). When $R > \hat{R}$ in fact, the Quasi-Discipline equilibrium ceases to attract, and Cooperation becomes the global attractor. Conversely, when condition (8) is satisfied, initial conditions matters. As in the other Trading Places regime, when the initial share of compliers and motivators are both large (resp., small), the system will eventually converge towards the Cooperation (resp., the Quasi-Discipline) equilibrium. The key difference between this and the other regimes investigated so far is that neither universal shirking nor universal compliance will occur when supervisors act as punishers. The reason behind this mechanism is that both types of workers have higher payoffs when they work with a punisher and a player of their own type than when they work with a punisher and a player of the other type. In other words, the incentives to comply are positive for the compliers who are matched with a fellow-complier but are negative for those who are matched with a defector. Formally, this is given by the fact that $w^H - E - Hp - K(1 - p) > w^L$ and that $w^H - E - Hp < (1 - p)w^H + pw^L$ are simultaneously satisfied in this regime. The mechanism works as follows. Differently from the Bad Company regime, the sense of procedural injustice is weak in this scenario (in particular, it is weaker than the wage premium), and each individual complier will stick to her strategy when she is matched with a defector. Conversely, as the detection probability is low, she will switch her strategy when she is matched with a complier. For the same reasons, each individual defector will keep shirking when she is matched with a complier (the detection probability is low) and will conversely switch strategy when she is matched with a defector (the sense of procedural justice is weak). Hence the populations’ composition in equilibrium. As usual, raising the premium for hard workers promotes universal compliance, while the effect of increasing control on workers’ behavior is ambiguous. In particular, it depends on the sign of $W - H$: if this is positive (resp., negative), the disciplining effect is larger (resp., smaller) than the crowding out effect, and

increasing monitoring stimulates (resp., hinders) workers cooperation⁹. This is in line with Frey (1993) and Chang and Lai (1999).

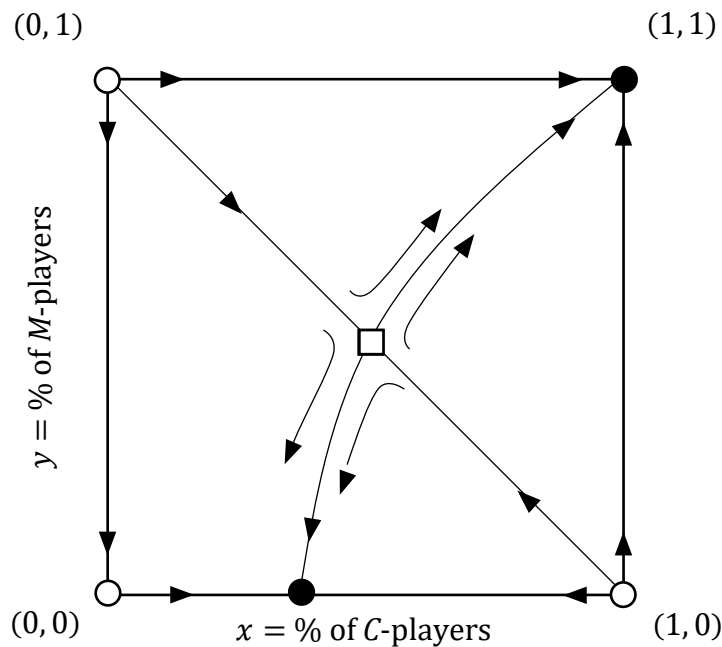


Fig. 4: Phase portraits of replicator dynamics (9) when condition (12) is satisfied (the “Trading Places II” regime). Filled dots represent attractors; empty dots represent repellers; empty squares represent saddle points and the two intersecting lines are the trajectories belonging to the stable and unstable branch of the internal saddle.

4.4 The “Disorder” regime

In the last configurations admitted for by dynamics (5), all four corners are saddle points, neither $(\hat{x}, 0)$ and $(\tilde{x}, 1)$ exist, and (\bar{x}, \bar{y}) is a source—see fig. 5. Differently from all other regimes investigated so far, it displays a perpetually oscillating behavior rather than convergence to a stable state¹⁰. The associated parametrization is given by:

⁹ Formally, this can be seen from the fact that $\partial(pW - E - Hp)/\partial p \stackrel{\geq}{\leq} 0$ if $W \stackrel{\geq}{\leq} H$, where $pW - E - Hp$ represents the payoff difference between complying and defecting when the supervisors if a punisher and the co-worker is a complier.

¹⁰ In a game where players are perfectly rational, (\bar{x}, \bar{y}) would correspond to a mixed-strategy Nash-equilibrium. When players have perfect foresight, they optimally play strategies C and M with equilibrium frequencies \bar{x} and \bar{y} respectively.

$$R^C > R^P \text{ and } q < \frac{E}{B} \text{ and } p > \frac{K + E - W}{K - H} \text{ and } p > \frac{E}{W - H} \quad (9)$$

The distinctive traits of this regime are the following. First, since motivators are not charismatic, they fail to engage compliers, and by doing so, they induce the latter to defect. When the supervisor is a motivator and $q < E/B$, in fact, the incentives to comply are negative, regardless of the coworkers' choice to comply or defect. Second, as monitoring efficiency is high, the incentives to comply are always positive when the supervisor is a punisher, regardless of the coworkers' choice to comply or defect. Formally, this is given by the fact that $w^H - E - Hp - K(1 - p) > w^L$ and $w^H - E - Hp > (1 - p)w^H + pw^L$ simultaneously hold when (9) is satisfied. Third, since workers are high skilled ($R^C > R^P$), motivators have higher payoff than punishers when they are matched with two compliers. Finally, recall that punishers have higher payoff than motivators when they are matched with two defectors. With these facts in mind, we can provide an explanation for the system's cyclical behavior.

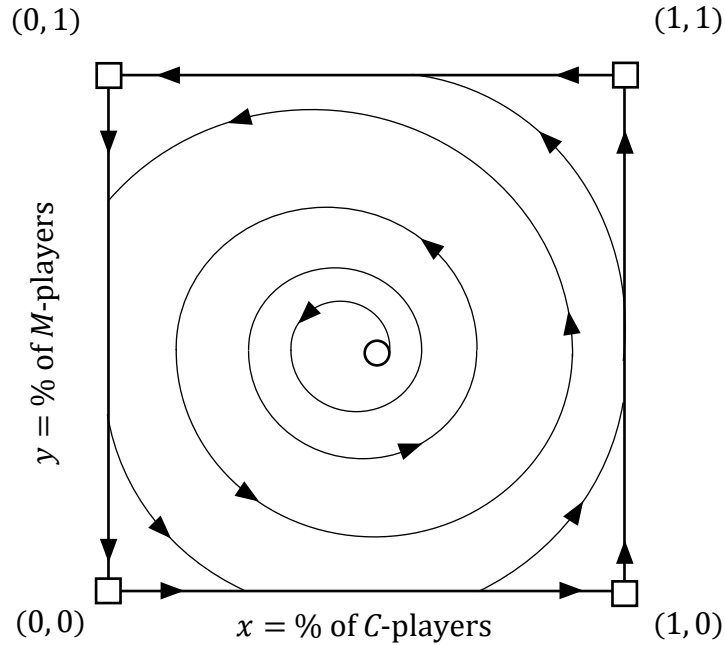


Fig. 5: Phase portraits of replicator dynamics (9) when condition (13) is satisfied (the “Disorder” regime). Filled dots represent attractors; empty dots represent repellers; empty squares represent saddle points and the spiral rotating counterclockwise represents exemplary trajectories which converge towards the boundaries of Q .

Consider an initial pairs of behavioral types in the neighborhood of the Cooperation equilibrium. Since motivators are not charismatic, compliers have an incentive to defect. As the rate of defectors in the overall population increases, so does the punishers', who take advantage of high-performing monitoring to self-protect from the workers' misbehavior. Since the punishment from defection is high regardless of team composition, defectors have now an incentive to comply. Yet, since $R^C > R^P$, workers are high-skilled, and while the rate of compliers increase, so does that of motivators. However, given the latter inability to motivate obedience, as the share of motivators increase, so does the defectors, and the cycle starts all over again.

This is where we appreciate the most the role of charisma in the leaders' strategic behavior. In our framework, workplace autonomy is profitable when workers are high skilled, but in order to get the most from this valuable human capital, supervisors must develop their coaching and relational skills. This is line with the idea recalled earlier whereby «managers have to switch from supervisory roles to acting as facilitators, as enablers and as people whose jobs are the development of people and their skills so that those people will be able to perform value-adding processes themselves» (Hammer and Champy, 1993: 82). The policy implication is that training programs for low and middle-level managers (who are normally in charge of superintending workers) should target the formation of these skills.

5. Welfare

In this section, we analyze the welfare properties of the game from the viewpoint of both workers and supervisors. By doing so, we investigate the possibility that the system may gravitate towards a Pareto-inferior situation because of adverse initial conditions or poor equilibrium selection (Carrera, 2018). According to Pareto's classic definition, a given state of the system is comparatively less efficient than another when moving from the former to the

latter the utility of at least one individual increases. According to the same definition, a given equilibrium is Pareto-optimal when moving from the former to any other reachable state the utility of all individuals decreases. To measure welfare, we compute the workers' average payoff from (3) when $x = 1$ and from (4) when $x = 0$. Similarly, we compute the supervisors' average payoff from (6) when $y = 1$ and from (7) when $y = 0$. Finally, when $x = \hat{x}$, $\Pi^C - \Pi^D = 0$, and the workers' average payoff can be equally computed from (3) or (4). Our findings are summarized in the following Proposition:

PROPOSITION 2—*In the Bad Company regime, the Discipline equilibrium (1,0) is always more efficient than the Conflict equilibrium (0,0) for both workers and supervisors. In the Trading Places I regimes, the Cooperation equilibrium (1,1) is always more efficient than the Conflict equilibrium (0,0) for both workers and supervisors. In the Trading Places II regime, the Cooperation equilibrium (1,1) is always more efficient than the Quasi-Discipline equilibrium (\hat{x} , 0) for both workers and supervisors. In addition, the Discipline, the Conflict and the Quasi-Discipline equilibrium are always inefficient for at least one population, while the Cooperation equilibrium is efficient for both workers and supervisors when $R^C > R^P$ and $q > \frac{E-W[1-(1-p)\hat{x}]}{B}$.*

From Proposition 2, two insights are worth noticing. The first is that shirking may have negative repercussions on workers' well-being even when it is sustained as an equilibrium strategy. The result is immediately verified by noticing that when two equilibria are locally attractive, the equilibrium with $x = 0$ (when nobody complies) or with $x = \hat{x}$ (when some comply and some defect) is always less efficient than the equilibrium with $x = 1$ (when

everybody complies). This assertion may still hold even when the latter equilibrium does not attract¹¹.

To the best of our knowledge, this is novel in the literature on workers' effort. That counterproductive work behaviors undermine organizational performance is trivial. Less intuitive, however, are the mechanisms whereby shirking becomes endemic even when it hinders workers' well-being. In the efficiency wage literature, workers decide to shirk if the utility from shirking is higher than the utility from complying (see for instance Shapiro and Stiglitz, 1984; Bowles, 1985). However, once that peer pressure and bounded rationality are allowed into the picture, individual choices become characterized by mass behavior. In our framework, this is represented by the fact that shirking inflicts a negative externality upon co-workers. Hence, when the share of defectors is large and the sense of procedural injustice is strong, each worker prefers receiving the low wage rather than feeling as the "village idiot" who complies when her coworker shirks.

In Chang and Lai (1999, 2002), workers derive psychological benefits from joining the high-effort group and they do so increasingly when the size of the latter augments. Therefore, they provide further support to the idea that shirking may become endemic when work morale is weak. Despite the similarity of these snowballing characteristics, however, their modelling strategy does not allow for the possibility that shirking may have negative repercussions on workers' well-being. Thus, the inclusion of team-working and procedural justice in the model significantly extends its predictive power. In addition, this "bad company" effect applies well to other contexts, such as child rearing or crime deterrence, where misbehaviors often become habitual because of negative peer pressure effects. According to both Graham and Weiner

¹¹ Indeed, when $E + H < W < (E + H)/p$, the Discipline equilibrium is repulsive but more efficient than the Conflict equilibrium. Similarly, when $(E - W)/B < q < E/B$, the Cooperation equilibrium is repulsive but more efficient than the Conflict equilibrium. In addition, when $[E - W(1 + p\hat{x} - \hat{x})]/B < q < E/B$, the Cooperation equilibrium is repulsive but more efficient than the Quasi-Discipline equilibrium. Finally, the Cooperation equilibrium it is always more efficient than the Discipline equilibrium, regardless of its stability properties.

(1996) and Wickert (2002), for instance, deviating pressure from peers is the main cause of adolescent misbehavior at school, while Rovken, Boer, Tolsma and Ruiters (2017) show that individuals who are emotionally close to a delinquent have a higher probability of becoming offenders themselves¹².

The second insight is that authoritarian modes of supervision may have negative effects on organizational performance even when they are sustained as an equilibrium strategy. The result is immediately verified by noticing that when two equilibria are locally attractive, the equilibrium with $y = 0$ (where all supervisors choose to punish) is always less efficient than the equilibrium with $y = 1$ (where all supervisors choose to motivate). This assertion may still hold even when the latter equilibrium does not attract¹³.

In this view, the diffusion of managerial cultures of a “command-and-control type” should be discouraged. Cross-referencing this idea with the results from Gordon (1994)—who shows that high supervisory intensity correlates positively with low wages/job security and little worker bargaining power—it may be argued that cooperative regimes of labor relations should be encouraged to facilitate the loosening of workplace control. Relying heavily on monitoring, indeed, seems to be the only mechanism for enhancing work effort when industrial conditions are poor.

The idea that authoritarian supervision may emerge as a second-best is even clearer in the Conflict equilibrium, where defecting is supported as an equilibrium strategy because of the combination of inefficient monitoring and low charisma. As both organizations and supervisors have negative payoffs at this state, our model highlights the risk ineffective authority. Loosely speaking, this seems consistent with the available evidence on the positive correlation between

¹² The literatures on negative peer pressure is vast. For further support to the statements in the next, see the references in Graham and Weiner (1996), Wickert (2002) and in Rovken, Boer, Tolmsa and Ruiters (2017).

¹³ Indeed, when $R^P \hat{x} + W[1 - \hat{x} - (1 - p)(\hat{x} - \hat{x}^2)] < R^C < R^P$, the Cooperation equilibrium is repulsive but more efficient for workers than the Quasi-Discipline equilibrium. In addition, the Cooperation equilibrium it is always more efficient than the Conflict equilibrium, regardless of its stability properties.

incarceration and recidivism (Cullen, Johnson and Nagin 2009), where punitive measures and negative peer pressure reinforce rather than inhibit the attitude to misbehave.

The natural corollary of these observations is that the Cooperation equilibrium is the only candidate to be simultaneously efficient for both populations (i.e., it is the Pareto-optimum of the system). Hence, the situation in which the Cooperation equilibrium is the global attractor is the most desirable for both workers and supervisors. From Proposition 1, it is easy to derive that (1, 1) may globally attract under two alternative parametrizations, the first of which is given by:

$$R^P < R^C < \hat{R} \text{ and } q > \frac{E}{B} \text{ and } p > \frac{K + E - W}{K - H} \text{ and } p > \frac{E}{W - H}$$

While the second is given by:

$$R^C > \hat{R} \text{ and } q > \frac{E}{B} \text{ and } p > \frac{K + E - W}{K - H}$$

As the Cooperation equilibrium is attractive when workers are high-skilled and supervisors are charismatic, organizations should design their training programs to facilitate supervisors to develop their motivational skills. In addition, they should also improve on-the-job training for their low-level workers and help the latter to develop their decisional skills.

From the above conditions, we also see that a given level of monitoring efficiency is required to sustain the choice of motivating as an equilibrium strategy. When this is not the case, employees may be given the temptation to shirk, and this, in turn, may force supervisors to resume a more authoritarian leadership style to self-protect from the risk of counterproductive work behaviors. This is where our model agrees with the standard predictions from efficiency wage theory, as it postulates that a given level of monitoring is required to obtain a «comprehensive truce in intraorganizational conflict» (Nelson and Winter, 1982: 110). However, our model suggests that charismatic leaders should avoid controlling

their employees and use the technologies at their disposal as a deterrence device. This is at odds with the spectacular increase in the capillarity of control recorded over the last years (Nussbaum, 1992; Tabak and Smith, 2005; Rosengren and Ottosson, 2016), which may have raised the sense of procedural justice in organizations (Alge, 2001) but also reduced the employee's morale and attitude towards self-control (Stanton, 2000).

6. Conclusion

In this paper, we have developed a simple evolutionary model to analyze the effect of different leadership styles on work morale and organizational performance. The key finding is that authoritative leadership improves both productivity and workers' well-being, provided that workers improve their decisional skills and supervisors foster their charisma. When either of these conditions is unmet, supervisors have no choice but securing obedience through coercion and control. In addition, when monitoring is inefficient, imposing authority may lead to a productivity-depressing situation where negative peer pressure encourage counterproductive work behaviors (i.e., the Conflict equilibrium attracts).

The key implication is that the use of authority in the superior/subordinate relationship is always a second-best. This is in line with the suggestions from leadership (Tannenbaum and Massarik, 1957; Bass, 1985; Yukl, 2010; Cunliffe, 2011) and parenting theory (Baumrind, 1971; Darling and Steinberg, 1993), which highlight the inefficiencies of overcontrolling subordinates. The idea of ineffective authority is loosely in line with the positive correlation between incarceration and recidivism (Cullen, Johnson and Nagin, 2009).

In addition, we have also highlighted the role of peers in affecting group behavior. In our framework, defectors inflict a negative externality upon compliers, who suffer a sense of procedural injustice when their loafing colleagues get away with their shirking. The introduction of procedural justice does not only add realism to the model, as it significantly

extends the variety of configurations admitted for by its dynamics. In particular, it allows for the emergence of what we called the “Bad Company” regime, where the system’s eventual configuration depends on initial conditions: when work morale is high, the presence of sufficient *C*-players prevent shirking to escalate; when work morale is low, shirking become widespread. This negative effect is consistent with similar mechanisms analyzed in other disciplines, such as criminology (Rovken, Boer, Tolmsa and Ruiters, 2017) and educational psychology (Graham and Weiner, 1996; Wickert, 2002), where the exposure to deviant behaviors is often identified as a predictor of misconducts.

The policy implications of our analysis unfold on two different levels. The former insists on the importance for organizations to fine-tune their training programs for both workers and supervisors. The latter, on the hand, is trickier and deals with the fact that “initial conditions matter”. A key limitation of evolutionary models like ours, in fact, is that the system’s dynamics may follow different paths according to the initial distribution of types across the interacting populations. How these distributions do initially form, however, is left out from the model. In our framework, expectations play a major role. Poor industrial relations may indeed provide both workers and supervisors with negative beliefs concerning the others’ behavior. When both parties perceive organizations as conflictual places, they may decide to “play tough”, with negative repercussions on the system’s evolution. In this framework, supporting long-lasting processes of cultural socialization is vital to help building a cooperative framework for workers and supervisors to interact in.

References

- Alge, B.J., 2001. "Effects of computer surveillance on perceptions of privacy and procedural justice". *Journal of Applied Psychology* 86, 797-804.
- Antoci A., A. Fiori Maccioni and P.L Sacco. 2017. "Self-protection, psychological externalities, and the social dynamics of fear". *Journal of Conflict Resolution* 61(2): 349-371.
- Baumrind D. 1971. "Current patterns of parental authority". *Developmental Psychology* 41(1): 1-103.
- Bass, B.M. 1985. *Leadership and performance beyond expectation*. New York, US: Free Press.
- Boltanski L. and E. Chiappello (2005) *The new spirit of capitalism*. New York: Verso.
- Bowles, S. 1985. "The production process in a competitive economy: Walrasian, neo-Hobbesian, and Marxian models". *American Economic Review* 75(1): 16 -36.
- Carrera, J.S. 2018. "Evolutionary dynamics of poverty traps". *Journal of Evolutionary Economics* 29(2), 611-630.
- Chang, J. and C. Lai. 1999. "Carrots or sticks? A social custom viewpoint on worker effort". *European Journal of Political Economy* 15:297-310
- Chang, J. and C. Lai. 2002. "Is the efficiency wage efficient? The social norm and organizational corruption". *The Journal of Scandinavian Economics* 104(1): 27-47.
- Coase, R.H. 1937. "The nature of the firm". *Economica* 4(16), 386-405.
- Cullen, F.T., C.L. Johnson and D.S. Nagin. 2009. "Imprisonment and reoffending". *Crime and Justice* 38(1): 115-200.
- Cunliffe, A., 2011. "Relational leadership". *Human Relations* 64(11), 1425 -1449.
- Darling N and L. Steinberg. 1993. "Parenting style as context: An integrative model". *Psychological Bulletin* 113(3): 487-496.
- Frey, B.S. 1993. "Does monitoring increase work effort? The rivalry with trust and loyalty". *Economic Inquiry* 31: 663-670.
- Garvin, D.A., Wagonfeld, A.B., Kind, L., 2013. *Google's Project Oxygen: do managers matter?* Harvard Business School Case, 313-110.
- Gordon, D. 1994. "A cross-national perspective on monitoring and supervision". *American Economic Review* 84(2): 375-379
- Graham, S. and B. Weiner. 1996. "Theories and principles of motivation". In Berliner D.C. and R.C. Calfee (Eds.), *Handbook of educational psychology* (pp. 63-84). New York, NY, US: Macmillan.
- Hammer, M. and J. Champy. 1993. *Reengineering the corporation: a manifesto for business revolution*. London, UK: Brearley.
- Jonker, L.B. and P.D. Taylor. 1978. "Evolutionarily stable strategies and game dynamics". *Mathematical Bioscience* 40: 145-156.
- Kamphorst, J.J.A. and O.H. Swank. 2016. "Don't demotivate, discriminate". *American Economic Journal: Microeconomics* 8(1): 140165.
- Kamphorst, J.J.A and O.H. Swank.2013. "When Galatea cares about her reputation: How having faith in your workers reduces their motivation to shine". *European Economic Review* 60: 91-104.
- Kristof-Brown A.L., R.D. Zimmerman and E.C. Johnson. 2005. "Consequences of individuals' fit at work: a meta-analysis of person-job, person-organization, person-group, and person-supervisor fit". *Personnel Psychology* 58(2): 281-342.
- Lazear, E., K. Shaw, K. and C. Stanton. 2015. "The value of bosses". *Journal of Labor Economics* 33(4): 823-862.
- Marx, K. 1982. *Capital. A critique of political economy. Volume I*. London, UK: Penguin Classic.
- Mintzberg H. 1971. "Managerial work: Analysis from observation". *Management Science* (18)2: 97-110.
- Nelson, R., and S. Winter. 1982. *An evolutionary theory of economic change*. Harvard, US: Belknap Press.
- Nussbaum, K., 1992. "Workers under surveillance". *Computerworld* 26(1):21.
- Rovken, J.J., G. de Boer, J. Tolsma and S. Ruiter. 2017. "How friends' involvement in crime affects the risk of offending and victimization". *European Journal of Criminology* 14(6): 697-719.
- Rosengren, C. and M. Ottosson. 2016. "Employee monitoring in a digital context". In: Daniels, J., K. Gregory and T. McMillan Cottom (eds) *Digital sociologies*. Bristol: Policy Press, 181-194.
- Shapiro, C. and J.E. Stiglitz. 1984. "Equilibrium unemployment as a worker discipline device". *American Economic Review* 74(3): 433-44.
- Stanton, J.M. 2000. "Traditional and electronic monitoring from an organizational justice perspective". *Journal of Business and Psychology* 15: 129-148.
- Swank, O.H. and B. Visser. 2007. "Motivating through delegating tasks or giving attention". *Journal of Law, Economics & Organization* 23(3): 732-742.
- Tabak, F., and W.P. Smith. 2005. "Privacy and electronic monitoring in the workplace: a model of managerial cognition and relational trust development". *Employees Responsibilities and Rights Journal* 17: 173-189
- Tannenbaum, R., and F. Massarik 1957. "Leadership: A frame of reference". *Management Science* 4(1): 1-19.
- Wickert, K. C. 2002. *Friends, cliques, and peer pressure: Be true to yourself*. Berkeley Heights: Enslow Publishers
- Yukl, G. 2010. *Leadership in organizations*. Upper Saddle River, US: Prentice-Hall.