WORKING PAPERS

Cooperation in Organizations through Self-Commitment Actions

Francesco Lancia
Alessia Russo

June 2016

Working Paper No: 1605

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

Cooperation in Organizations through Self-Commitment Actions*

Francesco Lancia Alessia Russo

June 21^{st}, 2016

Abstract

This paper studies how an organization might promote cooperation between its members when individual contributions to the organization's output are imperfectly observable. It considers an overlapping-generation game in which members with conflicting interests expend effort in pursuing outside tasks that are perfectly observable and privately beneficial in addition to the effort devoted to increasing the organization's output. We show that both the organization's expected output and members' well-being increase when the reward and punishment mechanism links the two types of effort. In the resulting equilibrium, privately beneficial efforts are at an inefficient level in order to signal members' willingness to cooperate. After extending the game to multiple generations, we apply it to the optimal tasks assignment along career paths in an organization.

JEL Classification: C73, D62, M54.

Keywords: Career Path, Common Good, Intergenerational Cooperation, Organizations, Public Perfect Equilibrium.

---

*Francesco Lancia, University of Vienna, email: francesco.lancia@univie.ac.at. Alessia Russo, University of Oslo, email: alessia.russo@econ.uio.no. A previous version of the paper was circulated with the title "Self-Commitment-Institutions and Cooperation in Overlapping Generations Games". We thank Daniel Garcia, Bård Harstad, David K. Levine, Anirban Mitra, Nicola Pavoni, Nicola Persico, Paolo Piacquadio, Debraj Ray, Karl Schlag, Paolo Siconolfi, and Timothy Worrall for helpful suggestions. We also benefitted from comments made by participants in the 2nd Personnel Economics and Public Finance conference in Ravello, the 11th SAET meeting in Faro, the 17th ISNIE conference in Firenze, and the seminars at the EIEF and the universities of Amsterdam, Bergen, Bologna, Firenze, Innsbruck, Modena, Napoli, Oslo, and Vienna. We also acknowledge the support of an EIEF research grant and the European Research Council under the EU's 7th Framework Programme, ERC GA no. 283236.

1 Introduction

Government agencies and private firms are classic examples of organizations which over time recruit junior members to replace senior ones who are retiring from the organization. When they join the organization, new members become recipients of, as well as contributors to, the organization's output. The success of these organizations therefore depends on their members' willingness to refrain from individually profitable actions and to work for the common good. The design of pecuniary incentive schemes constitutes the standard economic approach to this problem.[1] However, at a time when firms are cutting back on their financial incentive programs, as they are in most countries, non-pecuniary motivators may be an alternative way of building long-term engagement among members. The model presented here will show how the motivation to contribute to the organization's output can be aligned with social norms and the desire to build up an individual's reputation.

Ever since the seminal contribution of Hammond (1975), a vast literature has developed to analyze the sustainability of cooperation based on reputation in ongoing organizations. In this literature, cooperation can be sustained as an equilibrium by means of a chain of rewards and penalties spanning over generations (see, e.g., Cremer, 1986; Kreps, 1996).[2] This literature builds on the idea that collective decisions are perfectly observable and that members expend effort along one dimension. In the real world, however, contributions are largely unobservable. Cooperative relationships are thus typically plagued with free-rider problems and private and social optima do not coincide. Moreover, an organization's members can expend effort along several dimensions, some of which are only privately beneficial. For instance, tenured managers and employees within a firm, collaborators in a research team or physicians in countries with a dual health system can devote part of their physical and mental resources to private activities outside the organization, such as consulting activities for other organizations or training activities that provide them with additional qualifications or skills.

The effect of interactions among multiple actions and outcomes determines the power of incentives to achieve the common good. Such interactions depend on whether the actions are substitutes or complements in the member's payoff function. The research questions are therefore: When the details of cooperation between members are imperfectly observable, is it desirable for the organization to devise its reward and punishment mechanism so as to link privately beneficial effort to effort for the common good? Is this mechanism also to the benefit of the organization's members?

[1]See Besley and Ghatak (2015) for an overview of standard economic models of agency problems.

[2]Salant (1991), Kandori (1992), and Smith (1992) prove that if players overlap for a sufficiently long time and are patient, standard folk theorem results apply in an overlapping generations context.

To address these questions, we develop a model that considers an organization whose members belong to two different generations at any point in time: the junior member who joins the organization in the current period and the senior member who joined it in the previous period and who will exit before the next one. Within each period, members interact via a prisoner's dilemma type of game in which they simultaneously choose how much effort to devote to producing the organization's output. Total output depends on the sum of their efforts, which is subject to errors of observation. In addition to such collectively beneficial efforts, members can devote effort to tasks outside the organization, which are assumed to be perfectly observable and to benefit the agent alone. We refer to such activities as self-commitment actions when they exhibit complementarity or substitutability relations with the collectively beneficial effort. While the two generations have different time horizons, they share the same preferences and face the same strategic possibilities in each period. A member's preferences are fully characterized by a taste for both the organization's output and tasks outside the organization and a distaste for expending effort.

We focus on the Public Perfect Equilibrium (hereafter, PPE) and study the best rule for promoting cooperative behavior among the organization's members. We distinguish two regimes which differ in the history-contingent strategies available to the organization's members. The first is restricted, such that members do not condition their behavior on the history of self-commitment actions. The second is unrestricted, such that strategies are contingent on the entire history the members observe before playing. It follows that, while in the former regime, self-commitment actions are taken so as to maximize individual utility regardless of past history, in the latter self-commitment actions involve a strategic choice to manipulate the response of future generations. Comparing the best PPEs for each of the two regimes, two main predictions emerge, which involve the resolution of the tradeoff between efficiency and enforceability.

First, it is of value to the organization to employ a reward and punishment mechanism that specifies strategies contingent on the history of self-commitment actions. This is true even though self-commitment actions do not directly increase the organization's output. The intuition behind this result is as follows: Agents are always tempted to shirk since individual contributions are imperfectly observable. To deter deviation without compromising efficiency, the best strategy entails randomized punishment. As a result, members will forgive potential opportunistic behavior with a certain probability. Suppose now that members can choose a self-commitment action. Since the action is perfectly observable, the best strategy involves reversion to autarky with certainty if any agent deviates from the recommended level. This severe but credible deterrent induces all players to adopt the recommended level of self-commitment action in the first place.

Moreover, the effort must be expended at an inefficient level, which must be higher (lower) than that which members would choose if they were committed to cooperate in the case of complementarity (substitutability) with the collectively beneficial effort. By doing so, agents endogenously undermine their short-term gains from opportunistic behavior. All members then correctly internalize the changes in the individual incentive structure and accordingly adjust their retaliatory responses. Ultimately, the mechanism leads to higher forgiveness probability, while leaving enforceability unaltered and increasing the expected organization's output.

The second prediction is that it is of value to the organization's members to adopt strategies contingent on the history of self-commitment actions. The analysis leading to this result involves the comparison of the intertemporal utility that each member attains under the two alternative regimes. The difference between the two intertemporal utilities quantifies the strategic value of self-commitment actions from the members' perspective. We show that such a value is always positive regardless of the discount factor and monitoring technology. This occurs because the long-term gain of cooperation due to a higher equilibrium level of forgiveness always offsets the short-term cost born by the junior member in signaling his willingness to cooperate by choosing self-commitment actions.

As an additional prediction, we show that formal institutions, such as, for example, mandatory provision of activities for the common good or rules establishing a minimum amount of working time in the organization, undermine rather than promote the self-enforcement of cooperation. At first glance, this appears to contradict conventional wisdom. Although the imposition of a mandatory provision constraint reduces the gain from deviation for members and, in turn, increases the organization's output, it has a negative effect on the individual's incentives to comply with cooperation since it weakens the strategic role of self-commitment actions. Indeed, a tighter mandatory provision constraint has a twofold impact on the well-being of the organizations' members: First, it increases the marginal cost associated with the adoption of self-commitment actions at an inefficient level, and second, it lowers the marginal benefits from a higher level of forgiveness.

Although the baseline model is quite stylized, the strategic role of self-commitment actions survives even with the addition of more realistic features. For instance, the optimal level of self-commitment actions may vary with the size of the organization or the seniority of its members. We show how results generalize to an environment with more than two generations and interpret the results as the optimal task assignment along members' career paths within an organization. We demonstrate that requiring the youngest members to build up their reputations by means of outside activities is the best practice for the organization to adopt when the individual discount factor is sufficiently low. This is because the junior member's expected utility during his period as a senior member is small and therefore individual incentives to work to increase the organization's output are weak. Increasing members' efforts in activities outside the organization that are complements to inside activities, or reducing them when they are substitutes can boost motivation to cooperate between members.

The paper draws from the literature on strategic interactions with multiple actions initiated by Spence (1977) and Dixit (1980) and developed by Fudenberg and Tirole (1984) and Benoit and Krishna (1987), among others. The idea they put forward is that harsher punishments should improve incentives to cooperate by reducing the value of the players' outside option. This is the logic of deterrence, according to which an agent can invest preemptively in technology if such investments make it more costly for other agents to deviate. These models, however, assume that multiple actions are not interdependent in the payoff; in contrast, we focus on a supermodular game with strategic interactions among efforts.[3] This leads to a new mechanism involving self-commitment actions, which facilitates cooperation by endogenously undermining the short-run gain from deviation rather than by varying the retaliatory power of the punishment scheme.

The paper is therefore also related to the literature on issues linkage which uses supermodular games with multiple actions to study collusion in oligopolies with multi-market contracts (see, e.g., Bernheim and Whiston, 1998; Spagnolo, 1999; Limao, 2005). In these models, agents interact simultaneously on several issues, all characterized by the strategic structure of a prisoner's dilemma type of game. When issues are substitutes, linking them makes punishment harsher and deviation less worthwhile. The opposite holds when issues are complements. We depart from these models by embedding a self-interested action within a prisoner's dilemma. We find that enforcing self-commitment action at a privately inefficient level and linking them with collectively beneficial actions can increase cooperation whether decisions are substitutes or complements.[4]

Methodologically, the paper is related to the literature on repeated games with imperfect public monitoring, which studies optimal penal codes (see, e.g., Green and Porter, 1984; Radner, 1986; Abreu, Pearce, and Stacchetti, 1990; Abreu, Milgrom, and Pearce, 1991). Agents' inability to detect other players' opportunistic behavior with certainty

[3] See Topkis (1998) for a review of the theory of supermodularity and complementarity.

[4] Acemoglu and Jackson (2015) consider an overlapping generations game in order to study how a mechanism that specifies strategies contingent on the history of an action that is visible to all future agents can facilitate cooperation. Their model, however, differs from ours in several respects: First, observable actions are adopted by a leader who can influence the cooperative decisions of his followers. Second, the interactions between generations occur via a coordination game. Finally, there is no interdependence in the individual payoffs between observable and unobservable actions, which is the main feature of our analysis.

results in inefficient punishments. Fudenberg, Levine, and Maskin (1994) identify the conditions on the information structure which ensure that all feasible and individually rational payoffs can be supported in equilibrium. These results, however, do not hold when players are short-lived.[5] In this context, we show that cooperation increases when agents exert privately beneficial efforts in order to signal their willingness to cooperate.

Finally, the paper bears some relation to the multitasking and job design literature (see, e.g., Holmstrom and Milgrom, 1991 and 1994; Baker, Gibbons, and Murphy, 1994). When different tasks have heterogeneous measurable outcomes, these models show that offering stronger incentives to perform tasks that are more measurable distorts effort away from other tasks. The main implication in terms of job design and optimal allocation of tasks is that less discretion should be given to a manager in pursuing outside activities when inside effort is harder to measure. In contrast to that literature, we show that building up one's reputation through both inside and outside activities motivates members to increase their performance within the organization. Moreover, we show that the optimal tasks assignment that encourages members to pursue outside activities at an inefficient level is of value, not only to the organization but also to its members.

The paper proceeds as follows: Section 2 presents the model's setup and discusses benchmark results. Section 3 derives the worst and best PPE in each of the two regimes. Section 4 compares the two resulting upper bounds of PPE and discusses the effect of introducing mandatory provision constraints. Section 5 extends the basic model to an organization populated by more than two generations. Section 6 concludes. The proofs that are not in the text can be found in the Appendix.

2 The Setup

Time is discrete and indexed by $t = 0, 1, \dots$. The model consists of an ongoing organization with an overlapping generations demographic structure, whose members live for two periods and share the same discount factor $\delta \in (0, 1]$. Each generation is composed of one single member.[6] At each time t, a new member i enters the organization and is young (denoted by y) in the first period and old (denoted by o) in the second period.

[5] Bhaskar (1998) shows that in an overlapping generation context the existence of cooperative equilibria depends crucially on the observability of the entire history of play. In particular, no cooperation is the unique equilibrium in pure strategies when only a finite number of periods of past play can be observed. With no memory of previous play, Lagunoff and Matsui (2004) and Anderlini, Gerardi, and Lagunoff (2008) show that if agents are altruistic and can send messages, then a folk theorem holds.

[6] The assumption that there is one agent in each generation simplifies the analysis, but is not essential to the argument. The introduction of multiple agents within each generation would change the intratemporal incentive structure without modifying the intertemporal tradeoff that is the focus of the analysis.

2.1 Actions and Payoffs

The organization's members make choices in both periods of life and can exert effort along two dimensions: a privately beneficial effort and a collectively beneficial effort, which differ in both economic scope and observability. First, members decide on the privately beneficial effort, denoted by $b_t^i \in \mathbb{R}_+$, which is perfectly observable by both currently living agents and future generations. After this decision has been made for the current period, members choose a level of collectively beneficial effort, denoted by $a_t^i \in \{\underline{a}, \overline{a}\}$ with $0 \leq \underline{a} < \overline{a}$, which is not observable by other members.[7] This effort goes toward producing the organization's output $g_t = G\left(\sum_i a_t^i\right) + \epsilon$ where $G(\cdot)$ is an increasing function of individual effort and ϵ is an i.i.d. random variable with zero mean and cdf $F(\epsilon)$ independent of each a_t^i. Choosing the action $\overline{a}$ is interpreted as cooperation, whereas choosing the action $\underline{a}$ is interpreted as shirking. While the action a_t^i is not observable, the realized output g_t is perfectly observable by all generations. Exerting effort is costly, where the total cost $C(a_t^i, b_t^i)$ born by each member is strictly increasing in both types of effort and strictly convex in b_t^i (as described below).

Each member derives utility from the organization's output as well as from her privately beneficial actions. We assume quasi-linear preferences. Hence, the per-period utility of member i at time t can be written as:

$$u\left(g_t, a_t^i, b_t^i\right) := \lambda g_t + \theta b_t^i - C\left(a_t^i, b_t^i\right),$$

where $\lambda > 0$ captures the preference weight on the organization's output and $\theta > 0$ measures the marginal benefit of the privately beneficial action.

The organization's output is clearly maximized when both generations cooperate. However, we assume that at each time t member i prefers to shirk rather than cooperate for any level of collectively beneficial effort expended by the other member, denoted by a_t^{-i}, as well as for any level of privately beneficial effort b_t^i chosen by her. No cooperation therefore is the natural outcome of the organization in the absence of institutions that provide the necessary incentives to exert a high level of effort.

Assumption 1 *For any a_t^{-i} and b_t^i, the following relation holds:*

$$\lambda\left(G\left(\overline{a} + a_t^{-i}\right) - G\left(\underline{a} + a_t^{-i}\right)\right) < C\left(\overline{a}, b_t^i\right) - C\left(\underline{a}, b_t^i\right).$$

When privately and collectively beneficial efforts are interdependent in the individ-

[7]The sequential timing of the individual's decision making is not a critical assumption in the analysis when there are two generations in each period. However, it simplifies the analysis in Section 5, where the model is extended to more than two generations.

ual's utility function, a higher or lower level of effort along one dimension can have an impact on the individual's incentives to exert effort in the other dimension. Let $\Delta\left(b_t^i\right) := C_b\left(\overline{a}, b_t^i\right) - C_b\left(\underline{a}, b_t^i\right)$ measure the marginal impact of variation in b_t^i on the short-term gain of deviation from cooperative behavior for member i. Self-commitment actions are then defined as privately beneficial actions that satisfy the following property:

Definition 1 *A self-commitment action is any b such that $\Delta(b) \neq 0$.*

Thus, b is a self-commitment action as long as the per-period utility of each member exhibits interdependence between collectively and privately beneficial efforts. These two efforts can either be complements ($\Delta(b) < 0$), or substitutes ($\Delta(b) > 0$). In the former case, the incremental gain from choosing $\underline{a}$, i.e., the benefit from shirking, decreases as b increases. The opposite holds in the case of substitutability.[8]

These two types of effort are commonly observed in organizations. As an example of substitutability, employees can devote part of their physical and mental resources to private activities outside the organization, such as, for example, consulting activities for other firms or governments. This is effort that does not benefit the organization and substitutes for effort exerted to increase the organization's output. As an example of complementarity, employees can participate in training activities, which provide additional qualifications or skills that are privately beneficial, but may also reduce the marginal cost of effort devoted to inside activities. In the reminder of the paper, we analyze both cases in parallel.

2.2 Information Structure and Equilibrium Concept

The organization aims at maximizing realized output. However, it cannot rely on external enforcement to legally bind members to cooperate. Therefore, the only way to enforce cooperation is through repeated interaction. We follow the literature on self-enforcing agreements by casting the problem in an infinitely repeated-game setting. The Public Perfect Equilibrium (hereafter, PPE) serves as the equilibrium concept of the game between successive generations. A PPE induces a mapping from public history to the intertemporal utility of each member.[9]

We allow each member to condition her strategies on a public randomization $\phi_t \in [0,1]$, as is standard in the literature of repeated games with imperfect public monitoring (Mailath and Samuelson, 2006). We assume that such a device is drawn from

[8] A simple parameterization of the total cost of effort borne by each member is the quadratic form $C(a,b) = c\left[a^2 + 2kab + b^2\right]$, where $c > 0$, since this allows for the case of substitutability when $k > 0$ and the case of complementarity when $k < 0$.

[9] See Fudenberg and Tirole (1991) for a formal definition of PPE.

a uniform distribution at the beginning of each period and is perfectly observable by all members. The history of realizations of the organization's output, self-commitment actions, and public randomization up to time t are denoted by $h_g^t := (g_0, g_1, ..., g_{t-1})$, $h_b^t := \left(b_0^i, b_1^i, ..., b_{t-1}^i\right)_{i=y,o}$, and $h_\phi^t := (\phi_0, \phi_1, ..., \phi_t)$, respectively. Members' strategies map each history $h^t := \left(h_g^t, h_b^t, h_\phi^t\right)$ to the actions space $\{\underline{a}, \overline{a}\} \times \mathbb{R}_+$, where we refer to $g_s\left(h^t\right)$, $b_s\left(h^t\right)$, and $\phi_s\left(h^t\right)$ as the realizations g_s, b_s, and ϕ_s, respectively, in the public history h^t at time $s \leq t$ with $s \geq 0$. Since individual strategies are contingent on the full history of all preceding play, including the history of self-commitment actions, members strategically use such actions to manipulate future generations' responses. We will refer to this scenario as the unrestricted regime.

We wish to evaluate the strategic role played by self-commitment actions in promoting cooperation in an organization. To do so, we contrast the unrestricted regime and the restricted regime, in which members' strategies are contingent only on the history of the realization of both the organization's output and the public randomization. Formally, in this restricted regime, members' strategies map each history $\tilde{h}^t := \left(h_g^t, h_\phi^t\right)$ to the actions space $\{\underline{a}, \overline{a}\} \times \mathbb{R}_+$, where we refer to $g_s\left(\tilde{h}^t\right)$, $b_s\left(\tilde{h}^t\right)$, and $\phi_s\left(\tilde{h}^t\right)$ as the realizations g_s, b_s, and ϕ_s, respectively, in the public history $\tilde{h}^t$ at time $s \leq t$ with $s \geq 0$.

2.3 Benchmarks

Before embarking on the analysis of the strategic role played by self-commitment actions, it is worthwhile making a few observations. Under Assumption 1, the old member is always better off by shirking than cooperating, since she will be replaced by a successor in the following period. Absent altruistic motives regarding future generations, the old member receives no benefit from cooperation. We can then state the following preliminary result:

Proposition 1 *In any PPE, the efforts expended by the old member are* $(a_t^o, b_t^o) = (\underline{a}, b^{aut})$ *where* b^{aut} *solves* $\theta = C_b\left(\underline{a}, b^{aut}\right)$.

The old member does not contribute to the organization's output, while maximizing her own private benefit. Her presence, however, is of value to the organization since their future successors are able to form expectations of future payoffs, which they will benefit from when old. Depending on whether the successor cooperates or shirks, the expected utility of a member when old is equal to $\overline{\omega} := \lambda G\left(\overline{a} + \underline{a}\right) + \theta b^{aut} - C\left(\underline{a}, b^{aut}\right)$ or $\underline{\omega} := \lambda G\left(2\underline{a}\right) + \theta b^{aut} - C\left(\underline{a}, b^{aut}\right)$, respectively. Given the result of Proposition 1, we can rewrite the per-period utility of the young member simply as a function of her own effort levels, i.e., $u\left(a_t^y, b_t^y\right) := \lambda\left(G\left(a_t^y + \underline{a}\right) + \epsilon\right) + \theta b_t^y - C\left(a_t^y, b_t^y\right)$. Hereafter, unless otherwise specified, we shall omit the superscript y since the young member is the only one who might cooperate.

Having characterized the old member's behavior, we now turn to establishing two outcomes which will serve as benchmarks in the rest of the paper. First, consider the scenario in which young members also act non-cooperatively in each period. We refer to this scenario as generational autarky.

Proposition 2 *In a PPE that sustains generational autarky, i.e., $a_t^y = \underline{a}$, the self-commitment action expended by the young member is b^{aut} and her intertemporal utility is $v^{aut} := u\left(\underline{a}, b^{aut}\right) + \delta\underline{\omega}$.*

Second, suppose that the organization can rely on external enforcement. It would then be optimal for the organization to have both the young and old members cooperate, but the organization is not able to obtain cooperation from the old. Hence, the best outcome that the organization can hope for is to have the young cooperate. We refer to this scenario as constrained first-best.

Proposition 3 *In the case of constrained first-best, i.e., $a_t^y = \overline{a}$, the self-commitment action expended by the young member is b^*, which solves $\theta = C_b\left(\overline{a}, b^*\right)$, and her intertemporal utility is $v^* := u\left(\overline{a}, b^*\right) + \delta\overline{\omega}$.*

The following corollary shows that the level of self-commitment actions expended by the young in autarky is lower than that in the constrained first-best in the case of complementarity. The opposite holds true in the case of substitutability.

Corollary 1 *If $\Delta(b) < (>)\, 0$, then $b^* > (<)\, b^{aut}$.*

Proof. *(See Appendix).* ■

Cooperation is clearly desirable for the organization, since a high level of collectively beneficial effort maximizes the organization's output. We consider here the plausible case in which cooperation is desirable also for the young members, although at each time t non-cooperation is a strictly dominant strategy. This requires that $v^* > v^{aut}$, which is equivalent to the following restriction on the individual discount factor:

Assumption 2 $\delta > \underline{\delta} := \left(u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^*\right)\right) / \left(\overline{\omega} - \underline{\omega}\right)$.

To guarantee that the lower bound $\underline{\delta} \in [0, 1]$, it must be the case that $0 \leq u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^*\right) \leq \overline{\omega} - \underline{\omega}$.

3 Self-Enforcing Intergenerational Cooperation

We now turn to considering the case in which no external enforcement is available and the organization's members cannot commit to cooperation. We note that strategies contingent on past histories and enforced by reputation mechanisms allow for multiple equilibria. However, we confine our attention to the characterization of the best PPE, that is, the upper bound attainable by members without the help of external enforcement, which is determined by maximizing the intertemporal utility of each member. When the best PPE enforces the strongest cooperation among members, then it also delivers the organization's highest expected output. Depending on whether or not strategies are contingent on the history of self-commitment actions, two different upper bounds emerge. In order to determine what they are, we need to first characterize the worst PPE, which is common to both regimes.

3.1 The Worst Equilibrium

The worst PPE is the one that provides members the lowest intertemporal utility. This utility level coincides with v^{aut} according to the following argument: First, v^{aut} is sustainable. Indeed, if it is known that no one will ever cooperate, it is individually optimal not to cooperate. Second, there can be no lower equilibrium payoff, since each member is at the reservation utility.

Proposition 4 *The worst PPE is v^{aut}, which always exists.*

We will use the worst equilibrium as a threat to enforce better equilibria. If a pair (a_t^i, b_t^i) can be sustained in some PPE, then it can be sustained in a PPE where any deviation requires members to permanently revert to the worst PPE. When deviations occur from the perfectly observable action b_t^i, it is best to punish deviations with certainty. By contrast, when deviations occur from the imperfectly observable action a_t^i, cooperation must break down with some probability. Randomization of punishment under imperfect monitoring is required to minimize inefficient punishments, which occur on the equilibrium path. Hence, we can with no loss of generality focus on such simple trigger strategies to determine the upper bound of PPE intertemporal utilities both in the restricted and unrestricted regimes.[10]

[10] The equilibrium strategy is along the lines of Green and Porter (1984) who show that with imperfect monitoring firms can create collusive incentives by allowing price wars to break out with some probability. In the case of binary actions such a strategy also sustains the optimal equilibrium. See Abreu, Pearce, and Stacchetti (1986) for a characterization of optimal symmetric equilibria under imperfect monitoring.

3.2 The Best Equilibrium in the Restricted Regime

We start the analysis of the upper bound of PPE intertemporal utility, denoted by $\tilde{v}^e$, with the restricted case, in which members' strategies are contingent on the history $\tilde{h}^t$. This implies that, regardless of past history, the equilibrium self-commitment action is at the level that maximizes members' per-period utility, namely b^* when members cooperate and b^{aut} when members shirk.

Provided that $\tilde{v}^e \geq v^{aut}$, the best PPE is achieved by a simple trigger strategy, which can be represented by automata with two states: a cooperation state and an absorbing punishment state. Members start in the cooperation state. Regardless of the history h_b^t, the punishment state is activated with probability $1 - \phi$ when the organization's output in the previous periods is lower than a threshold level $\widehat{g}$. Namely, at $t = 0$:

(1) Start in the cooperation state by playing $\overline{a}$ and b^*;

At each $s \leq t$:

(2) If the organization's output is high, i.e., $g_s\left(\tilde{h}^t\right) > \widehat{g}$, then go back to (1);

(3) If the organization's output is low, i.e., $g_s\left(\tilde{h}^t\right) \leq \widehat{g}$, then go back to (1) with probability ϕ or revert permanently to generational autarky with probability $1 - \phi$.

Given the result of Proposition 1, a low output is realized with probability $p := F\left(\widehat{g} - G\left(2\underline{a}\right)\right)$ when the young shirk and $q := F\left(\widehat{g} - G\left(\overline{a} + \underline{a}\right)\right)$ when the young cooperate. We note that the monotone likelihood ratio property holds, i.e., $p > q$, and therefore the likelihood ratio is $L := p/q > 1$.[11]

Calculating the best PPE is equivalent to finding the largest ϕ for which the intertemporal utility of each young member is maximized and the self-enforcement constraint is satisfied. Formally:

$$\tilde{\phi}^e := \arg\max_{\phi} \ \tilde{v}\left(\phi\right), \tag{P1}$$

where

$$\tilde{v}\left(\phi\right) := \max_{b_t} u\left(\overline{a}, b_t\right) + \delta\left[\overline{\omega} - q\left(1 - \phi\right)\left(\overline{\omega} - \underline{\omega}\right)\right], \tag{1}$$

[11] Such a trigger strategy is not the only equilibrium strategy that can deliver the best PPE. However, we focus on it because of its simple structure. In this setting, the sufficient condition for the existence of a unique level $\widehat{g}$ is guaranteed by the monotone likelihood ratio property under which a tail test is the optimal statistical criterion for the players to adopt (Abreu, Pearce and Stacchetti, 1990). We could endogenize the level $\widehat{g}$ by maximizing L subject to the self-enforcement constraint. We note, however, that optimal monitoring possibilities are not necessarily accessible due to technological constraints. In the following analysis, we will consider p and q to be exogenous and perform comparative statics with respect to L.

subject to the constraint:

$$\tilde{v}(\phi) \geq \max_{b_t} u(\underline{a}, b_t) + \delta[\overline{\omega} - p(1-\phi)(\overline{\omega} - \underline{\omega})]. \tag{2}$$

Eq. (1) defines the intertemporal utility in the cooperation state. Inequality (2) describes the self-enforcement constraint regarding the cooperative action, which must be satisfied in order to discourage deviation from $\overline{a}$. Its right-hand side captures the intertemporal utility when agents deviate by choosing $\underline{a}$.[12] Manipulating Eqs. (1) and (2) yields:

$$\delta(1-\phi)(p-q)(\overline{\omega} - \underline{\omega}) \geq u\left(\underline{a}, b^{aut}\right) - u(\overline{a}, b^*). \tag{3}$$

The left-hand side of (3) is decreasing in ϕ and goes to zero when ϕ approaches one. This implies that cooperation cannot be sustained without punishment.[13] A lower ϕ, however, reduces members' intertemporal utility, as well as the organization's expected output since members remain in the cooperation state with a lower probability. This defines the tradeoff between *efficiency* and *enforceability*, which must be resolved. Solving the optimization program (P1) leads to the following proposition:

Proposition 5 (Necessary Condition) *Assume a PPE exists in which* $\tilde{v}^e \geq v^{aut}$ *and* $a_t = \overline{a}\ \forall t$. *Then,* $\phi \in [0, \overline{\phi}]$ *with* $\overline{\phi} \geq 0$ *and equal to*

$$\overline{\phi} := 1 - \frac{u(\underline{a}, b^{aut}) - u(\overline{a}, b^*)}{\delta(p-q)(\overline{\omega} - \underline{\omega})}. \tag{4}$$

In this case, the best PPE is unique and characterized by $\tilde{\phi}^e = \overline{\phi}$ *and*

$$\tilde{v}^e = u(\overline{a}, b^*) + \delta\overline{\omega} - \frac{u(\underline{a}, b^{aut}) - u(\overline{a}, b^*)}{L-1}. \tag{5}$$

If $\tilde{\phi}^e < 0$, *then the best such PPE yields* $\tilde{v}^e = v^{aut}$.

Proof. *(See Appendix).* ■

The best PPE is attained when $\tilde{\phi}^e = \overline{\phi}$. In fact, there can be no $\tilde{\phi}^e < \overline{\phi}$, since otherwise Eq. (1) might be further increased without violating constraint (2). The upper bound (5) corresponds to Abreu, Pearce, and Milgrom (1991)'s formula for the best pure trigger-strategy equilibrium payoff. It is equal to v^* minus the efficiency loss associated with the

[12]In equilibrium, it must also be true that members prefer to play $\underline{a}$ rather than cooperate in the punishment state. However, this condition is trivially satisfied under Assumption 2.

[13]Similar implications hold when δ approaches zero or when p tends to q. Indeed, in both cases, the temptation for an agent to deviate is so strong—either because of the agents' shortsightedness or because of the lack of informativeness of the performance variable regarding the collectively beneficial effort—that no degree of punishment has any deterrent power.

inefficient punishment, which occurs with some probability along the equilibrium path. For L approaching infinity, the efficiency loss associated with the inefficient punishment vanishes and $\tilde{v}^e$ tends to v^*.

It now remains to characterize the conditions under which $\tilde{\phi}^e \in [0,1]$. We note that $\tilde{\phi}^e \leq 1$ since $u(\underline{a}, b^{aut}) \geq u(\overline{a}, b^*)$. Then, the unique condition to be verified is $\tilde{\phi}^e \geq 0$. Using Eq. (4), the nonnegativity condition is satisfied when:

$$\delta \geq \tilde{\delta}^a := \frac{u(\underline{a}, b^{aut}) - u(\overline{a}, b^*)}{(p-q)(\overline{\omega} - \underline{\omega})}. \tag{6}$$

If $\delta < \tilde{\delta}^a$, then $\tilde{\phi}^e < 0$ and the unique sustainable equilibrium is generational autarky.

Proposition 6 (Sufficient Condition) *A threshold level $\tilde{\delta}^a \in [\underline{\delta}, 1]$ exists, so that $\tilde{v}^e \geq v^{aut}$ can be sustained as a PPE for any $\delta \geq \tilde{\delta}^a$.*

Proof. *(See Appendix).* ■

3.3 The Best Equilibrium in the Unrestricted Regime

We now turn to characterizing the upper bound of PPE intertemporal utility, denoted here by v^e, in the case where member's strategies are contingent on the history h^t. In this scenario, agents choose self-commitment actions while internalizing the strategic impact of such decisions on future members' responses.

Provided that $v^e \geq v^{aut}$, the best PPE is achieved by a trigger strategy that shares similar features with the strategy of Section 3.2. An additional punishment state, however, is considered here which dictates punishment with probability one when a member has deviated from the recommended level of self-commitment action b. It implies that the corresponding automata representation is characterized by three states: a cooperation state and two absorbing punishment states. Namely, at $t = 0$:

(1) Start in the cooperation state by playing $\overline{a}$ and b;

At each $s \leq t$:

(2) If the organization's output is high, i.e., $g_s(h^t) > \widehat{g}$, and $b_s(h^t) = b$, then go back to (1);

(3i) If the organization's output is low, i.e., $g_s(h^t) \leq \widehat{g}$, and $b_s(h^t) = b$, then go back to (1) with probability ϕ or revert permanently to generational autarky with probability $1 - \phi$;

(3ii) If $b_s(h^t) \neq b$ for any $g_s(h^t)$, then revert permanently to generational autarky with probability one.

The best PPE is achieved here by choosing the appropriate levels of ϕ and b such that the intertemporal utility of each young member is maximized and the self-enforcement constraints are satisfied. Formally:

$$(b^e, \phi^e) := \arg\max_{(b,\phi)} \; v(b, \phi), \tag{P2}$$

where

$$v(b, \phi) := u(\overline{a}, b) + \delta\left[\overline{\omega} - q(1-\phi)(\overline{\omega} - \underline{\omega})\right], \tag{7}$$

subject to the constraints:

$$v(b, \phi) \geq u(\underline{a}, b) + \delta\left[\overline{\omega} - p(1-\phi)(\overline{\omega} - \underline{\omega})\right], \tag{8}$$

$$v(b, \phi) \geq v^{aut}. \tag{9}$$

Eq. (7) defines the intertemporal utility in the cooperation state. In equilibrium, it must be that members prefer to cooperate as well as comply with b. Thus, two self-enforcement constraints must be satisfied: The first is inequality (8), which is the self-enforcement constraint regarding the cooperative action. It must be satisfied in order to discourage deviation from $\overline{a}$. Its right-hand side captures the intertemporal utility when agents deviate by choosing $\underline{a}$, albeit complying with the self-commitment action b. The second is inequality (9), which is the distinctive feature of (P2) and describes the self-enforcement constraint regarding the self-commitment action. It must be satisfied in order to discourage deviation from the recommended level b. In the case of deviation, punishment is triggered with certainty, given the observability of the actions. The individual's response to certain punishment is then to deviate from cooperation as well. Such an inequality acts as a participation constraint, which if not satisfied implies that agents have no incentive to join the organization.[14] Combining Eqs. (7) and (8) yields the following inequality:

$$\delta(1-\phi)(p-q)(\overline{\omega} - \underline{\omega}) \geq u(\underline{a}, b) - u(\overline{a}, b). \tag{10}$$

Eq. (10) highlights the tradeoff between efficiency and enforceability. Compared to inequality (3), the key difference is the right-hand side. If $(u(\underline{a}, b) - u(\overline{a}, b)) - (u(\underline{a}, b^{aut}) - u(\overline{a}, b^*))$ is larger than zero, then the tradeoff between efficiency and enforceability is even

[14]For the strategy to be an equilibrium, it must also be true that in the punishment states agents prefer to play $\underline{a}$ rather than cooperate, but this is trivially satisfied since punishment is not costly for the punisher, i.e., $u(\underline{a}, b) > u(\overline{a}, b)$.

exacerbated in this scenario as compared to the case where self-commitment actions are not taken strategically. Indeed, cooperation is enforced only if $\phi < \tilde{\phi}^e$, with the effect of depressing the individual's intertemporal utility. The reverse holds true if the difference is negative.

Having established the optimization program, we can now solve for the best PPE. For this purpose, let $\mathcal{S}$ be the set of pairs (b, ϕ) that satisfies constraints (8) and (9).

Proposition 7 (Necessary Condition) *Assume a PPE exists in which* $v^e \geq v^{aut}$ *and* $a_t = \overline{a}\ \forall t$. *Then,* $\mathcal{S} \neq \emptyset$, *i.e.,* $\phi \in \left[\underline{\phi}(b), \overline{\phi}(b)\right] \neq \emptyset\ \forall b$ *with*

$$\overline{\phi}(b) := 1 - \frac{u(\underline{a}, b) - u(\overline{a}, b)}{\delta(p-q)(\overline{\omega} - \underline{\omega})} \tag{11}$$

and

$$\underline{\phi}(b) := \frac{u(\underline{a}, b^{aut}) - u(\overline{a}, b)}{\delta q(\overline{\omega} - \underline{\omega})} - \frac{1-q}{q}. \tag{12}$$

In this case, the best PPE is unique and characterized by $\phi^e = \overline{\phi}(b^e)$ *where* b^e *solves* $u_b(\overline{a}, b^e) = \Delta(b^e)/(L-1)$ *and*

$$v^e = u(\overline{a}, b^e) + \delta\overline{\omega} - \frac{u(\underline{a}, b^e) - u(\overline{a}, b^e)}{L-1}. \tag{13}$$

If $\mathcal{S} = \emptyset$, *then the best such PPE yields* $v^e = v^{aut}$.

Proof. *(See Appendix).* ■

Proposition 7 predicts under which conditions $v^e \geq v^{aut}$. Suppose that a nonempty set $\mathcal{S}$ exists. Then, in equilibrium the following facts are necessarily true for any b: There can be no $\phi > \overline{\phi}(b)$; otherwise, inequality (8) would be violated. This would occur because of the temptation not to cooperate. Analogously, there can be no $\phi < \underline{\phi}(b)$; otherwise, inequality (9) would be violated. This would occur because of the temptation not to join the organization. From Eq. (7), moreover, we learned that the intertemporal utility is increasing in ϕ. We can then conclude that the best PPE must lie on the upper boundary of $\mathcal{S}$, namely $\phi = \overline{\phi}(b)$. Conditional on $\overline{\phi}(b)$ being feasible, the following corollary holds:

Corollary 2 *When privately and collectively beneficial efforts are complements (substitutes):*

(i) $\overline{\phi}(b)$ *is strictly increasing (decreasing) in* b;

(ii) $\forall b$, $\overline{\phi}(b)$ *is larger when* δ *and* p *are larger and* q *is smaller.*

Proof. *(See Appendix).* ■

Part (i) of Corollary 2 follows directly by differentiating $\overline{\phi}(b)$ with respect to b, which yields $\overline{\phi}_b = -\Delta(b) / (\delta(p-q)(\overline{\omega} - \underline{\omega}))$. Therefore, $\overline{\phi}_b > 0$ when $\Delta(b) < 0$ and $\overline{\phi}_b < 0$ otherwise. When the two actions are complements, the members' inference that a deviation from cooperation occurred in the past is strengthened as the level of b decreases. This leads to a lower level of forgiveness probability and in turn to a lower expected utility when old. Indeed, once agents have internalized that a larger b reduces the individual marginal gain from defection and have observed that previous members have invested a high level of effort in a self-commitment action, they consistently believe that the realization of a low output is to be attributed primarily to a negative shock that is not under the members' control. The reverse argument applies when efforts are substitutes. Part (ii) of Corollary 2 states that $\overline{\phi}(b)$ is lower and, in turn, cooperation is harder to sustain when agents are less patient (i.e., δ is smaller) or monitoring is weaker (i.e., p approaches q). In these cases, members are more tempted to cheat. For any b, generations must therefore punish with higher probability in order to discourage deviation from cooperative behavior.

Corollary 3 *When privately and collectively beneficial efforts are complements (substitutes):*

(i) $b^e > (<)\, b^*$;

(ii) $|b^* - b^e|$ *is decreasing with* L.

Proof. *(See Appendix).* ■

Corollary 3 provides two additional insights: Part (i) implies that b^e always differs from the constrained first-best level, regardless of whether the collectively and privately beneficial efforts are complements or substitutes. This occurs because an inefficient level of self-commitment action may improve the continuation utility by reducing the inefficiency loss associated with the punishment, i.e., it enlarges $\overline{\phi}(b)$, even though it reduces the young's per-period utility in the cooperation state, i.e., it lowers $u(a, b)$. This is true when $b^e > b^*$ in the case of complementarity and $b^e < b^*$ in the case of substitutability. Part (ii) predicts that a better monitoring technology weakens the strategic role of self-commitment actions, namely the difference between b^e and b^* shrinks as L increases. Intuitively, when the organization's output is a good indicator of the individual contributions, a member's motive to exert inefficient privately beneficial effort vanishes.

Figure 1 illustrates these results for the case of $\Delta(b) > 0$ (left-side panel) and $\Delta(b) < 0$ (right-side panel). In both panels, the shaded area depicts $\mathcal{S}$, whose upper boundary is downward-sloping when efforts are substitutes and upward-sloping when they are complements. The graph also plots the young's indifference curves of utility (7). The intertemporal utility increases as ϕ increases and $|b^* - b^e|$ shrinks. Therefore, it is maximized at

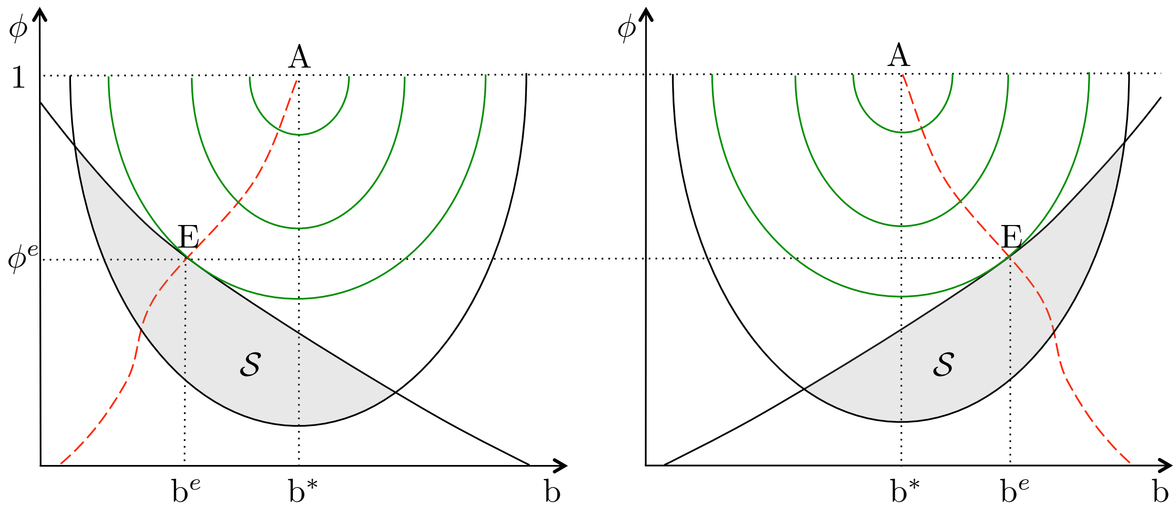

Figure 1: The Best PPE in the Unrestricted Regime

point A. However, such a maximum is not enforceable since constraint (8) is violated at this point. The highest intertemporal utility compatible with constraints (8) and (9) is achieved at the tangent point of the indifference curve with $\overline{\phi}(b)$, namely, at point E. At such a point, $b^e \neq b^*$and is located to the left of it when $\Delta(b) > 0$ and to the right of it otherwise. Finally, the graph depicts the locus (b^e, ϕ^e) resulting from variations in L and represented by the dotted red line, which approaches point A as L tends to infinity.

It now remains to determine when $(b^e, \phi^e) \in \mathcal{S} \neq \emptyset$. This is guaranteed whenever the conditions $\phi^e \geq 0$ and $\phi^e \geq \underline{\phi}(b^e)$ simultaneously hold.[15] Using Eqs. (11) and (12), this is the case when:

$$\delta \geq \delta^a := \frac{u(\underline{a}, b^e) - u(\overline{a}, b^e)}{(p-q)(\overline{\omega} - \underline{\omega})} \tag{14}$$

and

$$\delta \geq \delta^b := \frac{u(\underline{a}, b^{aut}) - u(\overline{a}, b^e)}{\overline{\omega} - \underline{\omega}} + \frac{q}{p-q} \frac{u(\underline{a}, b^e) - u(\overline{a}, b^e)}{\overline{\omega} - \underline{\omega}}. \tag{15}$$

If $\delta < \delta^a$, then $\phi^e < 0$. This implies that constraint (8) is violated and therefore $v^e = v^{aut}$. Similarly, if $\delta < \delta^b$, then $\phi^e < \underline{\phi}(b^e)$. This implies that members have no incentive to join the organization when a level of self-commitment action equal to b^e is required of them. Once again, the best PPE intertemporal utility is v^{aut}. Depending on the max $\{\delta^a, \delta^b\}$, either constraint (8) or constraint (9) can bind first. In both cases, we can establish the following result:

Proposition 8 (Sufficient Condition) *A threshold level* max $\{\delta^a, \delta^b\} \in [\underline{\delta}, 1]$ *exists,*

[15] Conditional on being positive, we note that $\phi^e < 1$. Moreover, conditional on $\phi^e > \underline{\phi}(b^e)$, $\underline{\phi}(b^e) < 1$. Thus, the only two conditions to verify in order to guarantee non-emptyness of $\mathcal{S}$ in equilibrium are the nonnegativity of ϕ^e and the relation $\phi^e \geq \underline{\phi}(b^e)$. See the Appendix for a formal discussion.

so that $v^e \geq v^{aut}$ can be sustained as a PPE for any $\delta \geq \max\left\{\delta^a, \delta^b\right\}$.

Proof. *(See Appendix).* ■

4 Strategic Value of Self-Commitment Actions

We have so far determined the best PPE in the two alternative regimes. While the equilibrium forgiveness probability is affected by varying the self-commitment action in the unrestricted regime, it is immune to strategic manipulation in the restricted one. Therefore, the following question naturally arises: Which of the two regimes provides the most value to the organization as well as to its members? The comparison between $\tilde{\phi}^e$ and ϕ^e directly implies the following prediction:

Proposition 9 *Strategic self-commitment actions are always of value to the organization, i.e., $\tilde{\phi}^e < \phi^e$.*

Proof. *(See Appendix).* ■

The result of Proposition 9 is intuitive. In equilibrium, requiring members to build up their reputation by means of outside tasks minimizes the members' gain from shirking. Moreover, inspecting $\overline{\phi}(b)$ as reported in Eq. (11) makes clear the lower are the short-term benefits from deviation, the higher will be the probability that agents forgive. A lower retaliation probability also implies that cooperation is more likely in the long run and that the organization's expected output is larger. For this reason, devising a reward and punishment mechanism so as to link privately beneficial effort to effort for the common good is the best practice for the organization to adopt.

Unlike the organization, members face the following intertemporal tradeoff: a short-term loss due to the reduction in per-period utility when young, as implied by Corollary 3, versus a long-term gain associated with a higher equilibrium forgiveness probability and, in turn, a higher expected utility when old, as stated in Proposition 9. To evaluate which of the two effects prevails, we introduce the function $W : [\underline{\delta}, 1] \to \mathbb{R}$, defined as the difference between the payoffs (13) and (5), i.e., $W := v^e - \tilde{v}^e$. This function captures the members' net surplus generated by enforcing strategies contingent on the history of self-commitment actions as compared to strategies that are not contingent on this type of information. Now consider a discount factor sufficiently large that both v^e and $\tilde{v}^e$ are strictly larger than v^{aut}, i.e., $\delta > \max\left\{\tilde{\delta}^a, \max\left\{\delta^a, \delta^b\right\}\right\}$. The members' net surplus can then be written as follows:

$$W = \overbrace{-\left(u\left(\overline{a}, b^*\right) - u\left(\overline{a}, b^e\right)\right)}^{\text{Short-Term Loss}} + \frac{1}{L-1}\overbrace{\left(\left(u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^*\right)\right) - \left(u\left(\underline{a}, b^e\right) - u\left(\overline{a}, b^e\right)\right)\right)}^{\text{Long-Term Gain}} \quad (16)$$

The following proposition shows how self-commitment actions can resolve the trade-off between efficiency and enforceability in organizations when the extent of cooperation is imperfectly observable. In particular, it shows that it is possible to increase members' well-being (point (i)) and at the same time expand the possibilities for enforcing cooperation (point (ii)).

Proposition 10 *Strategic self-commitment actions are always of value to the organization's members:* (*i*) *for any* $\delta > \max\left\{\tilde{\delta}^a, \max\left\{\delta^a, \delta^b\right\}\right\}$, $W > 0$; (*ii*) $\max\left\{\delta^a, \delta^b\right\} < \tilde{\delta}^a$.

Proof. *(See Appendix).* ■

The result of part (i) fundamentally hinges on the imperfection of the monitoring technology. To grasp the intuition, reformulate Eq. (16) as $(1/(L-1))(u(\underline{a}, b^{aut}) - u(\underline{a}, b^e)) - L/(L-1)(u(\overline{a}, b^*) - u(\overline{a}, b^e))$. Recall that $b^* > (<)\, b^{aut}$ in the case of substitutability (complementarity). Together with the result stated in Corollary 2, this implies that $|b^e - b^{aut}|$ is strictly larger than $|b^e - b^*|$ for any L. Therefore, the variation of the first component, $(1/(L-1))(u(\underline{a}, b^{aut}) - u(\underline{a}, b^e))$, due to changes in L, is of first-order impact compared to the variation of the second component, $L/(L-1)(u(\overline{a}, b^*) - u(\overline{a}, b^e))$, followed by a commensurate change in the likelihood ratio. This ultimately implies that surplus (16) is larger than zero for any degree of monitoring.

Part (ii) of Proposition 10 leverages this result to establish that when self-commitment actions are taken strategically, it is also possible to expand the scope of cooperation. When agents are so impatient that cooperation cannot be sustained in the organization, i.e., $\delta < \tilde{\delta}^a$, strategies contingent on the history of self-commitment actions can resolve such an enforceability issue. This is because cooperation can be sustained also when $\delta \geq \max\left\{\delta^a, \delta^b\right\}$, where $\max\left\{\delta^a, \delta^b\right\}$ proves to be smaller than $\tilde{\delta}^a$. It turns out that W is positive also for $\delta \in [\max\left\{\delta^a, \delta^b\right\}, \tilde{\delta}^a)$, whereas it is equal to zero for $\delta < \max\left\{\delta^a, \delta^b\right\}$.

Figure 2 illustrates the results of Proposition 10 for the cases $\delta^b > \delta^a$ (left-side panel) and $\delta^b < \delta^a$ (right-side panel). The upper part of the figure plots the best PPE intertemporal utilities in each of the two regimes. The solid-black line denotes v^e, while the dashed-red line denotes $\tilde{v}^e$. Both intertemporal utilities are stuck at v^{aut} insofar as $\delta \in \left(\underline{\delta}, \max\left\{\delta^a, \delta^b\right\}\right)$ in the case of v^e and $\delta \in \left(\underline{\delta}, \tilde{\delta}^a\right)$ in the case of $\tilde{v}^e$. After that, they may jump to a higher value and monotonically converge to their maximum sustainable values when δ approaches one.[16] The lower part of the graph plots the members' net surplus.

[16]A formal discussion of the discontinuity in the map W is provided in the proof of Proposition 6 and 8 in the Appendix.

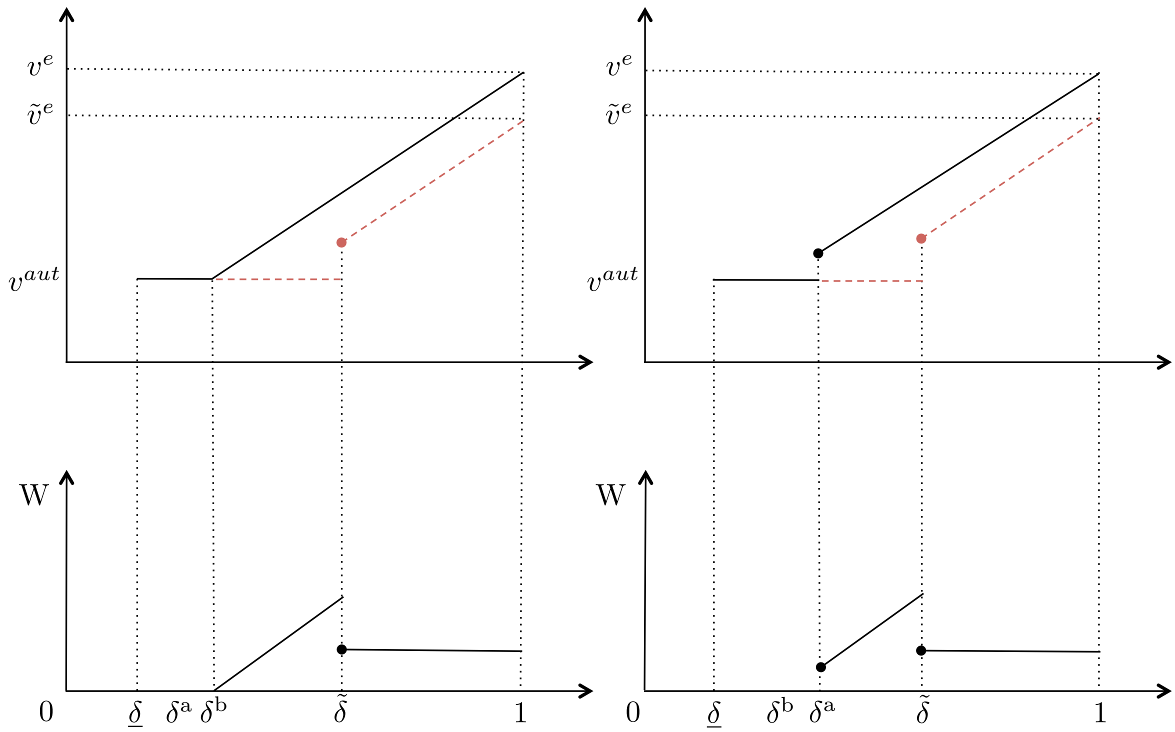

Figure 2: Members' Net Surplus

4.1 The Effect of Mandatory Provision

The previous analysis showed that the strategic interaction between collectively and privately beneficial efforts is of value to the organization, as well as to its members, and this is true whether the two efforts are complements or substitutes. A question that naturally arises is whether the introduction of a mandatory provision constraint on the cooperative action can further improve the organization's output and the members' net surplus. These types of constraints are common in organizations. For example, workers in a firm's worker can be required to carry out a minimum number of tasks or to work a minimum number of hours.

To introduce mandatory provision, we slightly modify the basic model by considering a compulsory minimum collectively beneficial effort $a^{\min} > \underline{a}$. The larger $a^{\min}$, the tighter is the mandatory provision constraint. The imposition of such a constraint clearly increases the organization's expected output, since it reduces the individual's gain from deviation and in turn increases the probability of forgiveness. However, it has a negative effect on the members' net surplus, as stated in the following proposition:

Proposition 11 *When $a^{\min}$ is increased, (i) $|b^e - b^*|$ shrinks and (ii) W is reduced.*

Proof. *(See Appendix).* ■

The introduction of a minimum provision constraint by means of an increase in $a^{\min}$ crowds out the strategic role played by self-commitment actions. Indeed, as $a^{\min}$ increases, the short-term gain from opportunistic behavior decreases. This reduces the marginal benefits generated by the enforcement of self-commitment action at the inefficient level, while increasing the marginal cost of its implementation.[17]

This result has interesting implications in terms of how formal institutions should be modelled in the presence of informal rules within an organization. While the imposition of a minimal provision constraint on individual contributions eliminates the possibility of the equilibrium with low cooperation, it also reduces the value of cooperation that can be sustained when the organization employs a reward and punishment mechanism that links privately beneficial effort to effort for the common good.

5 Multiple Generations

The framework of the model rests on the assumption that only two generations inhabit the organization at any time t. To make the model more realistic we now extend to more than two generations. The model can then be applied to the description of the optimal tasks assignment along career paths in an organization. We show that it may be optimal to require members who are more reluctant to cooperate to pursue outside tasks at an inefficient level so as to motivate them to try harder in tasks that are profitable for the organization.

It is straightforward to extend the basic model to an organization whose members live for three periods: young, middle-aged, and old. Define m to denote the middle-aged member.[18] As before, we focus on the characterization of the best PPE, which requires the largest number of generations to choose $\overline{a}$. This can be accomplished by using the trigger strategy described in Section 3.3 and appropriately choosing a level of randomization ϕ. Given that individual contributions are unobservable, the identity of agents who deviate from cooperation cannot be inferred by observing the organization's realized output. Hence, punishment in equilibrium must be triggered with the same probability in all generations. Clearly, the result of Proposition 1 still holds. Conditional on whether other

[17] Bernheim and Whinston (1998) obtain results with a similar flavor for non-intergenerational contracting problems. They show that when complete contracts are impossible and voluntary cooperation is needed in the dimensions where the contract is incomplete, it might be optimal to leave some dimensions out of the contract in order to increase the incentive to cooperate in the dimensions that cannot be included. Rangel (2003) shows that in an overlapping generation framework the introduction of a minimum provision of pensions reduces the sustainable maximum level of investment in future generations. Therefore, as in our setting, minimum provision constraints undermine cooperation if cooperation is already sustained by means of informal institutions.

[18] The probabilities of generating a bad signal in the case of individual deviation generalizes to $p := F\left(\widehat{g} - G\left(3\underline{a}\right)\right)$ and in the case of cooperation to $q := F\left(\widehat{g} - G\left(2\overline{a} + \underline{a}\right)\right)$.

members cooperate or shirk, the expected utility of a member when old is then equal to $\overline{\omega} := \lambda G(2\overline{a} + \underline{a}) + \theta b^{aut} - C(\underline{a}, b^{aut})$ or $\underline{\omega} := \lambda G(3\underline{a}) + \theta b^{aut} - C(\underline{a}, b^{aut})$, respectively. For notational purposes, let $u(a^i, b^i) := \lambda(G(a^i + \overline{a} + \underline{a}) + \epsilon) + \theta b^i - C(a^i, b^i)$ denote the per-period utility of member i when the old member shirks and the other members cooperate. The intertemporal utility in the cooperation state for the middle-aged and young members can therefore be written respectively as:

$$v^m(\phi, b^m) := u(\overline{a}, b^m) + \delta(\bar{\omega} - q(1-\phi)(\overline{\omega} - \underline{\omega})) \tag{17}$$

and

$$v^y(\phi, b^y, b^m) := u(\overline{a}, b^y) + \delta\left(v^m(\phi, b^m) - q(1-\phi)\left(v^m(\phi, b^m) - \left(u^{aut} + \delta\underline{\omega}\right)\right)\right) \tag{18}$$

where $u^{aut} := \lambda(G(3\underline{a}) + \epsilon) + \theta b^{aut} - C(\underline{a}, b^{aut})$ is the per-period utility in generational autarky. For the candidate strategy to be an equilibrium, it must be the case that in the cooperation state middle-aged members prefer $\overline{a}$ and b^m over deviating to $\underline{a}$ or b^{aut}, which yields, respectively:

$$v^m(\phi, b^m) \geq u(\underline{a}, b^m) + \delta(\overline{\omega} - p(1-\phi)(\overline{\omega} - \underline{\omega}))$$

or

$$v^m(\phi, b^m) \geq u^{aut} + \delta\underline{\omega}.$$

In addition, it must be the case that in the cooperation state young members prefer $\overline{a}$ and b^y over deviating to $\underline{a}$ or b^{aut}, which yields, respectively:

$$v^y(\phi, b^y, b^m) \geq u(\underline{a}, b^y) + \delta\left(v^m(\phi, b^m) - p(1-\phi)\left(v^m(\phi, b^m) - \left(u^{aut} + \delta\underline{\omega}\right)\right)\right)$$

or

$$v^y(\phi, b^y, b^m) \geq u^{aut} + \delta\left(u^{aut} + \delta\underline{\omega}\right).$$

Given the sequential timing of an individual's decision making and the perfect observability of b^i, deviating from the recommended level of self-commitment action triggers reversion to autarky starting from the current period. In the previous sections, we saw that depending on the model's fundamentals, either the self-enforcement constraint regarding the cooperative decision or the self-enforcement constraint regarding the self-commitment action can bind first. In this section, we simplify the exposition by confining our attention to the case in which compliance with a self-commitment action is necessarily guaranteed insofar as compliance with cooperation is enforced. Hence, we can simply focus on the constraints on the cooperative decision for middle-aged and young members, which can be written respectively as:

$$\phi \leq 1 - \frac{u\left(\underline{a}, b^{m}\right) - u\left(\overline{a}, b^{m}\right)}{\delta\left(p-q\right)\left(\overline{\omega}-\underline{\omega}\right)} \tag{19}$$

and

$$v^{m}\left(\phi, b^{m}\right) \geq \frac{u\left(\underline{a}, b^{y}\right) - u\left(\overline{a}, b^{y}\right)}{\delta\left(p-q\right)\left(1-\phi\right)} + \left(u^{aut} + \delta\underline{\omega}\right). \tag{20}$$

As in Eq. (8), middle-aged members are tempted to deviate when the forgiveness probability is overly high, while young members are tempted to deviate when the intertemporal utility of the middle-aged is overly low. We note that an increase in ϕ has a twofold impact on constraint (20): on the one hand, it increases $v^{m}\left(\phi, b^{m}\right)$, thereby reducing the young member's temptation to deviate, while on the other hand, it increases the right-hand side of (20), thereby making deviation more profitable for the young.

We now have all the ingredients in order to characterize the best PPE. Since Eq. (18) is increasing in $v^{m}\left(\phi, b^{m}\right)$, which in turn is increasing in ϕ, the best PPE involves the selection of the maximum level of ϕ compatible with inequalities (19) and (20) being satisfied. The enforceability constraint that binds first identifies the generation most reluctant to cooperate. On the one hand, the young may have little incentive to shirk since they have a long career ahead of them and in turn will benefit from a long period of cooperation; on the other hand, they may also be highly tempted to deviate since they face a higher probability of being trapped in the punishment state during their lifetime and therefore attain with a lower probability the stake $\overline{\omega}$ when old. Two effect that prevails depends on δ. The following proposition characterizes the equilibrium life-cycle profile of self-commitment actions when the best PPE is enforced:

Proposition 12 *The best PPE satisfying $a^{i} = \overline{a}$ for each i is characterized by:*

(i) $b^{y,e} = b^{*}$, $b^{m,e} > (<)\, b^{*}$ *when* $\Delta\left(b^{m}\right) < (>)\, 0$, *and* $b^{o,e} = b^{aut}$ *when the middle-aged member is the most reluctant;*

(ii) $b^{y,e} > (<)\, b^{*}$ *when* $\Delta\left(b^{y}\right) < (>)\, 0$, $b^{m,e} = b^{*}$, *and* $b^{o,e} = b^{aut}$ *when the young member is the most reluctant.*

Proof. *(See Appendix).* ■

When the discount factor is sufficiently large, the young always comply with cooperation whenever the middle-aged also do. It follows that members in the early stage of their careers are highly motivated to build long-term engagement in the organization. In that case, the use of non-pecuniary motivators in the form of strategic self-commitment actions is not optimal. In contrast, middle-aged members must signal their willingness to cooperate by partially giving up short-term private gains. Finally, old members maximize their private benefits with no concern for the common good. As a result, part (i) of Proposition 12 predicts that the life-cycle profile of self-commitment actions is either hump-shaped when privately and collectively beneficial efforts are complements or inverted hump-shaped when they are substitutes.

A different pattern emerges when the discount factor is sufficiently small that the young have the strongest incentives to deviate. In this scenario, an inefficient level of self-commitment effort is optimal in order to motivate the young to contribute a high level of effort to the organization's output. In contrast, the middle-aged maximize their private benefits without compromising their individual incentives to cooperate. The resulting life-cycle profile of self-commitment actions is then either monotonically decreasing or monotonically increasing depending on whether the two efforts are complements or substitutes, respectively.

The results suggest that organizations seeking to maximize output should assign different tasks to members at different stages of their career. When the discount factor is small and agents have low expectations of eventually being promoted to senior positions, it is optimal to allow them to pursue outside tasks and thus build up their reputations early in their career in order to increase their motivation. In contrast, when the discount factor is large, incentives provided by means of optimal task assignment in outside activities should be designed so as to motivate more experienced members to exert more effort in tasks that are of value to the organization.

6 Conclusions

The model attempts to capture the role played by privately beneficial action in promoting cooperation in an organization when individual contributions to the common good are imperfectly observable. Two fundamental features of the model drive the results: The first is the payoff interdependence among actions, which must exist in order for the organization to determine, at the margin, the power of incentives to work for the common good. The second is the repeated interaction among members. Agents adopt an inefficient level of self-commitment action only if they anticipate that such a choice will positively affect the incentives to cooperate of future players and thus raise their continuation value.

The two aforementioned features fundamentally distinguish the model from existing theories of commitment. Commitment devices may improve cooperation in a static setting without requiring payoff interdependence. However, the potential gains from cooperation are unlikely to be realized for the following three reasons: First, maintaining cooperation may require mutual pledges of commitment through written contracts, which are often difficult to enforce. Second, commitment is profitable only if it is not overly costly and not activated along the equilibrium path. In the presence of monitoring imperfection, however, agents must periodically incur the cost that they have committed to. Finally, the commitment solution requires the existence of an outside entity to construct commitment spaces for the players, such that cooperation is attainable as an equilibrium outcome in the restricted action space. In contrast to existing theories of commitment, our mechanism accomplishes cooperation in a plausible fashion. Self-commitment actions are self-enforcing. They are not conceived as a punishment, but rather as an equilibrium strategy, which we show to always have value to both an organization and its members in the presence of monitoring imperfection. Moreover, since self-commitment actions are part of the players' strategy space, they render unnecessary the existence of a third party outside of the game who is able to construct commitment spaces for the players.

The idea of self-commitment actions was used to study the sustainability of cooperation in ongoing organizations, such as government agencies and private firms. Clearly, there are many other interesting applications that fit into our setting. An example is the study of cooperation in religious organizations, which are ongoing and exist in order to achieve the common good of their members (see, e.g., Iannaccone, 1992; Levy and Razin, 2012). In this context, the model might be useful in highlighting the role of religious practices which, like self-commitment actions, can dictate the behavior of members. Another prominent example might be the study of socially responsible practices in corporations, which can be viewed as an optimal communal response in an uncertain environment, such as a volatile financial market. If such socially responsible practices reduce gains from unobservable deviation, then they can also foster goodwill and trust among shareholders (see, e.g., Baron, 2001 and 2010).

7 Appendix

Proof of Proposition 1. First consider the case of complementary efforts, i.e., $C_b\left(\overline{a}, b_t^i\right) - C_b\left(\underline{a}, b_t^i\right) < 0\ \forall b_t^i$. If $b_t^i = b^*$, then $\theta = C_b\left(\overline{a}, b^*\right) < C_b\left(\underline{a}, b^*\right)$ given the convexity of the function $C\left(\cdot\right)$. It follows that b^{aut} cannot be larger than b^*, otherwise $C_b\left(\underline{a}, b^{aut}\right) > C_b\left(\underline{a}, b^*\right) > \theta$ which would contradict $C_b\left(\underline{a}, b^{aut}\right) = \theta$. An analogous argument holds true for the case of substitute efforts. ■

Proof of Proposition 5. For a fixed level of $\widehat{g}$, there are only two signals available. Therefore, it is not possible to reward the young when the old are punished. It is then optimal to lower the punishment level to the point at which agents are indifferent between complying and shirking from cooperation. These observations lead to Eqs. (1) and (2) in the text. Using these equations, we obtain that (2) holds when $\phi \leq \overline{\phi}$, with $\overline{\phi}$ as reported in (4). The equilibrium forgiveness probability is then $\tilde{\phi}^e = \overline{\phi}$. Inserting Eq. (4) into (1), we obtain that the upper bound of PPE is equal to Eq. (5). If $\tilde{\phi}^e < 0$, then Eq. (2) is not satisfied, which implies that the upper bound is v^{aut}. ■

Proof of Proposition 6. Condition (6) is feasible when $\tilde{\delta}^a \in (\underline{\delta}, 1]$. It is straightforward to show that $\tilde{\delta}^a > \underline{\delta}$. Furthermore, $\tilde{\delta}^a \leq 1$ when

$$q \leq p - \frac{u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^*\right)}{\left(\overline{\omega} - \underline{\omega}\right)}. \tag{21}$$

Condition (21) delimits a non-empty region in the space (p, q). Hence, we conclude that there always exists a feasible $\tilde{\delta}^a$, such that: (i) if $\delta > \tilde{\delta}^a$, then $\tilde{\phi}^e > 0$ and $\tilde{v}^e > v^{aut}$; (ii) if $\delta = \tilde{\delta}^a$, then $\tilde{\phi}^e = 0$ and $\tilde{v}^e > (=)\, v^{aut}$, depending on whether $p < (=)\, 1$. ■

Proof of Proposition 7. We use the same considerations as in the proof of Proposition 5, except that members' strategies are contingent on the full history, which includes the history of self-commitment actions. As before, it is optimal to lower the punishment level to the point at which agents are indifferent between complying and shirking from cooperation. Given that self-commitment actions are perfectly observable and taken before the cooperative decision, it is clearly optimal to revert to generational autarky, i.e., the worst PPE, with certainty after a deviation has been observed. Furthermore, since we are looking at the best PPE, the equilibrium b is the one that maximizes the intertemporal utility of the young. These observations lead to Eqs. (7), (8), and (9) in the text. Using these equations, we obtain that (8) holds when $\phi \leq \overline{\phi}\left(b\right)$, whereas (9) holds when $\phi \geq \underline{\phi}\left(b\right)$, with $\overline{\phi}\left(b\right)$ and $\underline{\phi}\left(b\right)$ as reported in (11) and (12), respectively. The best PPE is attained by setting $\phi = \overline{\phi}\left(b\right)$ whenever $\overline{\phi}\left(b\right) \geq \underline{\phi}\left(b\right)$. Inserting Eq. (11) into (7), we obtain that the upper bound of PPE is equal to Eq. (13). If $\overline{\phi}\left(b\right) < \underline{\phi}\left(b\right)$, then there exists no feasible ϕ that simultaneously satisfies Eqs. (8) and (9), which implies that the upper bound is v^{aut}. ■

Proof of Corollary 2**.** The proof of part (i) appears in the text. To prove part (ii), differentiate $\overline{\phi}(b) := 1 - \frac{u(\underline{a},b)-u(\overline{a},b)}{\delta(p-q)(\overline{\omega}-\underline{\omega})}$ with respect to δ, p, and q. It follows that $\overline{\phi}_\delta = \frac{u(\underline{a},b)-u(\overline{a},b)}{\delta^2(p-q)(\overline{\omega}-\underline{\omega})} > 0$, $\overline{\phi}_p = \frac{u(\underline{a},b)-u(\overline{a},b)}{\delta(p-q)^2(\overline{\omega}-\underline{\omega})} > 0$, and $\overline{\phi}_q = -\frac{u(\underline{a},b)-u(\overline{a},b)}{\delta(p-q)^2(\overline{\omega}-\underline{\omega})} < 0$. ■

Proof of Corollary 3**.** Start by proving part (i). The equilibrium level of self-commitment action solves $\theta - C_b(\overline{a}, b^e) - (1/(L-1))\Delta(b^e) = 0$. Hence, $\theta - C_b(\overline{a}, b^e) > (<) 0$ when $\Delta(b^e) > (<) 0$. Since b^* is the level that solves $\theta - C_b(\overline{a}, b^*) = 0$, it follows that $b^e < (>) b^*$ given the convexity of the function $C(\cdot)$. We now prove part (ii). Using the implicit function theorem, we get $b_L^e = -\frac{\left(1/(L-1)^2\right)\Delta(b^e)}{-C_{bb}(\overline{a},b^e)-(1/(L-1))\Delta_b(b^e)}$, where the denominator is the second-order condition with respect to b, which must be negative. Hence, $b_L^e > (<) 0$ insofar as $\Delta(b^e) > (<) 0$. The result for part (i) and the derivative $b_L^* = 0$ guarantee that $|b^* - b^e|$ decreases with L. ■

Proof of Proposition 8**.** For $\phi^e \geq 0$ to be true, condition (14) must hold. This condition is feasible when $\delta^a \leq 1$, which implies that:

$$q \leq p - \frac{u(\underline{a}, b^e) - u(\overline{a}, b^e)}{\overline{\omega} - \underline{\omega}}. \tag{22}$$

It must also be the case that condition (14) satisfies $\delta^a > \underline{\delta}$. Then, it follows that:

$$q > p - \frac{u(\underline{a}, b^e) - u(\overline{a}, b^e)}{u(\underline{a}, b^{aut}) - u(\overline{a}, b^*)}. \tag{23}$$

Moreover, $\phi^e \geq \underline{\phi}(b^e)$ implies that condition (15) must hold. In this case, a feasible δ satisfying this condition exists when $\delta^b \leq 1$, i.e.,

$$q \leq \frac{(\overline{\omega} - \underline{\omega}) - (u(\underline{a}, b^{aut}) - u(\overline{a}, b^e))}{(\overline{\omega} - \underline{\omega}) - (u(\underline{a}, b^{aut}) - u(\underline{a}, b^e))} p. \tag{24}$$

It is straightforward to show that $\delta^b > \underline{\delta}$. For the existence of a feasible $\delta > \max\left\{\delta^a, \delta^b\right\}$, therefore, inequalities (22), (23), and (24) must be simultaneously satisfied. From part (i) of Corollary 2 and of Corollary 3, we learned that $\phi^e > \overline{\phi}(b^*)$, which implies that the inequality $u(\underline{a}, b^e) - u(\overline{a}, b^e) < u(\underline{a}, b^*) - u(\overline{a}, b^*)$ holds. Since $u(\underline{a}, b^*) < u(\underline{a}, b^{aut})$, the relation $u(\underline{a}, b^e) - u(\overline{a}, b^e) < u(\underline{a}, b^{aut}) - u(\overline{a}, b^*)$ is necessarily true, which guarantees that the term $\frac{u(\underline{a},b^e)-u(\overline{a},b^e)}{u(\underline{a},b^{aut})-u(\overline{a},b^*)}$ in condition (23) is smaller than one. From Assumption 2, moreover, we know that $u(\underline{a}, b^{aut}) - u(\overline{a}, b^*) < \overline{\omega} - \underline{\omega}$, which implies that the term $\frac{u(\underline{a},b^e)-u(\overline{a},b^e)}{\overline{\omega}-\underline{\omega}}$ in condition (22) is smaller than $\frac{u(\underline{a},b^e)-u(\overline{a},b^e)}{u(\underline{a},b^{aut})-u(\overline{a},b^*)}$. Then, conditions (22) and (23) delimit a non-empty region in the space (p, q). Hence, we can conclude that there always exists a feasible threshold for the discount factor equal to $\max\left\{\delta^a, \delta^b\right\}$, such that if $\delta > \max\left\{\delta^a, \delta^b\right\}$, then the set $\mathcal{S} \neq \emptyset$ and $v^e > v^{aut}$.

Finally, we note that $\delta^a \geq (<)\, \delta^b$ if and only if

$$\frac{u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^e\right)}{u\left(\underline{a}, b^e\right) - u\left(\overline{a}, b^e\right)} \leq (>)\, \frac{1-q}{p-q}.$$

If $\delta = \delta^a = \max\left\{\delta^a, \delta^b\right\}$, then $\phi^e = 0$. In this case, $v^e > (=)\, v^{aut}$ depending on whether $p < (=)\, 1$. In the alternative scenario with $\delta = \delta^b = \max\left\{\delta^a, \delta^b\right\}$, the equilibrium outcome prescribes $\underline{\phi}\left(b^e\right) = \phi^e = 0$ and $v^e = v^{aut}$ for any level of p. ■

Proof of Proposition 9**.** From the proof of Proposition 8, we learned that the relation $u\left(\underline{a}, b^e\right) - u\left(\overline{a}, b^e\right) < u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^*\right)$ is necessarily true. This directly implies that $\phi^e > \tilde{\phi}^e$. ■

Proof of Proposition 10**.** We first prove part (i). For any $\delta > \max\left\{\tilde{\delta}^a, \max\left\{\delta^a, \delta^b\right\}\right\}$, we can rewrite Eq. (16) as $W = (1/\left(L-1\right))(u\left(\underline{a}, b^{aut}\right) - u\left(\underline{a}, b^e\right)) - L/\left(L-1\right)\left(u\left(\overline{a}, b^*\right) - u\left(\overline{a}, b^e\right)\right)$. Differentiating W with respect to L and applying the envelope condition yields:

$$W_L = -\frac{1}{\left(L-1\right)^2}\left(u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^*\right) - \left(u\left(\underline{a}, b^e\right) - u\left(\overline{a}, b^e\right)\right)\right).$$

From Proposition 9, we obtained that $u\left(\underline{a}, b^e\right) - u\left(\overline{a}, b^e\right) < u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^*\right)$. Thus, $W_L < 0$ for any L. Moreover, W tends to zero for L approaching infinity. Therefore, it has to be that W approaches zero from above. This implies that W is positive for any level of L when $\delta > \max\left\{\tilde{\delta}^a, \max\left\{\delta^a, \delta^b\right\}\right\}$. We now prove part (ii). Consider the following two cases: First, let $\delta^a = \max\left\{\delta^a, \delta^b\right\}$. By Proposition 9, $u\left(\underline{a}, b^e\right) - u\left(\overline{a}, b^e\right) < u\left(\underline{a}, b^{aut}\right) - u\left(\overline{a}, b^*\right)$. Then, comparing Eq. (6) to Eq. (14), it is straightforward to show that $\tilde{\delta}^a > \delta^a$. Second, let $\delta^b = \max\left\{\delta^a, \delta^b\right\}$. If $\tilde{\delta}^a < \delta^b$, then $v^e = v^{aut}$ and $\widetilde{v}^e > v^e = v^{aut}$ for $\delta = \delta^b$. Using a continuity argument, $v^e > v^{aut}$ and $\widetilde{v}^e > v^{aut}$ for any $\epsilon > 0$ such that $\tilde{\delta}^a + \epsilon > \delta^b$, if $\delta = \tilde{\delta}^a + \epsilon$, which contradicts the result of part (i), namely that $W > 0$ for any $\delta > \max\left\{\tilde{\delta}^a, \max\left\{\delta^a, \delta^b\right\}\right\}$. ■

Proof of Proposition 11**.** We first prove part (i). Consider $a^{\min} \in (\underline{a}, \overline{a}]$. Then, the per-period utility is $u\left(a^{\min}, b\right)$ when the young do not cooperate. The equilibrium self-commitment decision is obtained solving the first-order condition $u_b\left(\overline{a}, b^e\right) - (1/\left(L-1\right))\left(u_b\left(a^{\min}, b^e\right) - u_b\left(\overline{a}, b^e\right)\right) = 0$. By the implicit function theorem:

$$b^e_{a^{\min}} = -\frac{-\left(1/\left(L-1\right)\right) u_{a^{\min}b}\left(a^{\min}, b^e\right)}{u_{bb}\left(\overline{a}, b^e\right) - \left(1/\left(L-1\right)\right)\left(u_{bb}\left(a^{\min}, b^e\right) - u_{bb}\left(\overline{a}, b^e\right)\right)},$$

where the denominator is the second-order condition, which is clearly satisfied in equilibrium. The sign of $u_{a^{\min}b}$ is the opposite to that of $\Delta(b^e)$. Since $b^e > (<)\, b^*$ when $\Delta(b^e) < (>)\, 0$ and $b^*_{a^{\min}} = 0$, we obtain that $|b^e - b^*|$ is decreasing in $a^{\min}$. We now prove part (ii). Differentiating W with respect to $a^{\min}$ and applying the envelope condition yields:

$$W_{a^{\min}} = \frac{1}{L-1}\left(\left(u_{a^{\min}}\left(a^{\min}, b^{aut}\right) - u_{a^{\min}}\left(a^{\min}, b^e\right)\right) - u_b\left(a^{\min}, b^e\right)\frac{\partial b^e}{\partial a^{\min}}\right),$$
$$\simeq -u_{a^{\min}b}\left(a^{\min}, b\right)\left[b^e - b^{aut}\right] - u_b\left(a^{\min}, b^e\right)\frac{\partial b^e}{\partial a^{\min}}.$$

Using the results from Corollary 1, part (i) of Corollary 3, and part (i) of Proposition 11, we have that $u_{a^{\min}b} > 0$, $b^{aut} < b^* < b^e$, $u_b\left(a^{\min}, b^e\right) < 0$, and $\frac{\partial b^e}{\partial a^{\min}} < 0$ when $\Delta(b) < 0$, which implies that $W_{a^{\min}} < 0$. Equivalently, $u_{a^{\min}b} < 0$, $b^e < b^* < b^{aut}$, $u_b\left(a^{\min}, b^e\right) > 0$, and $\frac{\partial b^e}{\partial a^{\min}} > 0$ when $\Delta(b) > 0$, which once again implies that $W_{a^{\min}} < 0$. ■

Proof of Proposition 12**.** As pointed out in the text, both the young and middle-aged members comply in the cooperation state if and only if constraints (19) and (20) are simultaneously satisfied. Figure 3 provides a representation of the set of sustainable equilibria satisfying the self-enforcement constraints, where $v^y(\phi, b^y, b^m)$ is the indifference curve of the young member, for any given level of b^y and b^m. The first panel illustrates part (i) of the proposition where constraint (19) is tighter than constraint (20). Part (ii) of the proposition is illustrated in the second panel. The shaded area represents the set of allocations (ϕ, v^m) that satisfies the two self-enforcement constraints simultaneously.

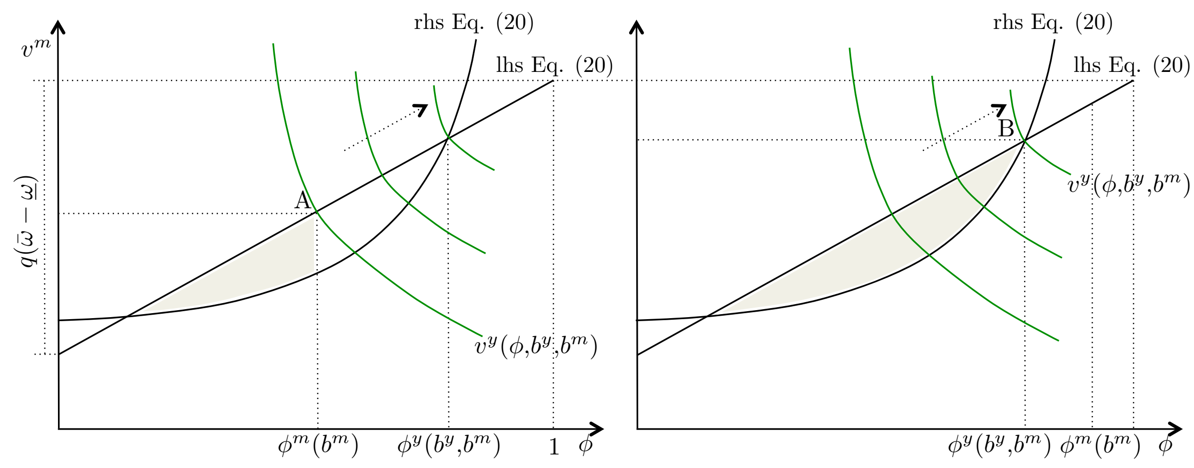

Figure 3: Self-Commitment Actions with Multiple Generations

We Start by proving part (i). Consider the case in which constraint (19) is binding whereas constraint (20) is slack. We determine ex-post the conditions under which this case holds true. The best PPE requires that:

$$\phi = \phi^m(b^m) := 1 - \frac{u(\underline{a}, b^m) - u(\overline{a}, b^m)}{\delta(p-q)(\overline{\omega} - \underline{\omega})} \tag{25}$$

Then, the best payoff of the middle-aged member is obtained by maximizing $v^m(\phi, b^m)$ with respect to b^m subject to Eq. (25). Hence, the following condition must hold:

$$u_{b^m}(\overline{a}, b^{m,e}) - \frac{q}{p-q}\Delta(b^{m,e}) = 0 \tag{26}$$

Given the concavity of $u(\overline{a}, b^m)$ in b^m, we get that $b^{m,e} < (>)\, b^*$ when $\Delta(b^m) > (<)\, 0$. The best payoff of the young is obtained by maximizing $v^y(\phi, b^y, b^m)$ with respect to b^y subject to Eq. (25) and $b^m = b^{m,e}$. This implies that $u_{b^y}(\overline{a}, b^{y,e}) = 0$ and thus, $b^{y,e} = b^*$. Finally, we determine the condition under which the constraint (19) is binding whereas the constraint (20) is slack. This requires the discount factor to be sufficiently large, such that:

$$\delta > \delta^m := \frac{q}{p-q}\frac{u(\underline{a}, b^{m,e}) - u(\overline{a}, b^{m,e})}{\overline{\omega} - \underline{\omega}} - \frac{u(\overline{a}, b^{m,e}) - u^{aut}}{\overline{\omega} - \underline{\omega}} + \frac{u(\underline{a}, b^{y,e}) - u(\overline{a}, b^{y,e})}{u(\underline{a}, b^{m,e}) - u(\overline{a}, b^{m,e})}$$

For reasonable values of p and q, it is straightforward to verify that there always exists a feasible δ satisfying the above condition.

Now consider part (ii) in which constraint (19) is slack and constraint (20) is binding. In this context, the best PPE requires that $\phi = \phi^y(b^y, b^m)$, where $\phi^y(b^y, b^m)$ is the maximum solution of the equation:

$$\phi = 1 - \frac{u(\underline{a}, b^y) - u(\overline{a}, b^y)}{\delta(p-q)(v^m(\phi, b^m) - (u^{aut} + \delta\underline{\omega}))} \tag{27}$$

The right-hand side of (27) is increasing and concave in ϕ. The implicit function theorem yields:

$$\phi^y_{b^y} = -\frac{\frac{\Delta(b^y)}{\delta(p-q)(v^m(\phi,b^m)-(u^{aut}+\delta\underline{\omega}))}}{1 - \frac{(u(\underline{a},b^y)-u(\overline{a},b^y))v^m_\phi}{\delta(p-q)(v^m(\phi,b^m)-(u^{aut}+\delta\underline{\omega}))^2}} \tag{28}$$

and

$$\phi^y_{b^m} = \frac{\frac{(u(\underline{a},b^y)-u(\overline{a},b^y))u_{b^m}(\overline{a},b^m)}{\delta(p-q)(v^m(\phi,b^m)-(u^{aut}+\delta\underline{\omega}))^2}}{1 - \frac{(u(\underline{a},b^y)-u(\overline{a},b^y))v^m_\phi}{\delta(p-q)(v^m(\phi,b^m)-(u^{aut}+\delta\underline{\omega}))^2}} \tag{29}$$

We note that $\frac{(u(\underline{a},b^y)-u(\overline{a},b^y))v^m_\phi}{\delta(p-q)(v^m(\phi,b^m)-(u^{aut}+\delta\underline{\omega}))^2} \leq 1$ at $\phi = \phi^y(b^y, b^m)$. Then, it follows that $\phi^y_{b^y} < (>)\, 0$ when $\Delta(b^y) > (<)\, 0$. Moreover, $\phi^y_{b^m} \geq (<)\, 0$ if and only if $u_{b^m}(\overline{a}, b^m) \geq (<)\, 0$. We now determine the equilibrium levels $b^{m,e}$ and $b^{y,e}$. Differentiating Eq. (18) with respect to b^m yields the following first-order condition:

$$u_{b^m}\left(\overline{a}, b^{m,e}\right) = -\left(\begin{array}{c} \frac{q}{1-q(1-\phi^y(\cdot))}\left(v^m\left(\phi^y\left(\cdot\right), b^{m,e}\right) - \left(u^{aut}+\delta\underline{\omega}\right)\right) \\ +v^m_{\phi^y}\left(\phi^y\left(\cdot\right), b^{m,e}\right) \end{array}\right)\phi^y_{b^m} \tag{30}$$

where the term in the brackets is positive. Then, there can be no $u_{b^m}\left(\overline{a}, b^{m,e}\right) > (<)\, 0$, otherwise we would have $\phi^y_{b^m} < (>)\, 0$, which would contradict the conditions from Eq. (29). Hence, the only solution that satisfies Eq. (30) is $u_{b^m}\left(\overline{a}, b^{m,e}\right) = 0$ and $\phi^y_{b^m} = 0$, which implies that $b^{m,e} = b^*$. Differentiating (18) with respect to b^y, we obtain the following first-order condition:

$$u_{b^y}\left(\overline{a}, b^{y,e}\right) = -\delta\left(1 - q\left(1-\phi^y\left(\cdot\right)\right)\right)\left(\begin{array}{c} \frac{q}{(1-q(1-\phi^y(\cdot)))}\left(v^m\left(\phi^y\left(\cdot\right), b^{m,y}\right) - \left(u^{aut}+\delta\underline{\omega}\right)\right) \\ +v^m_{\phi^y}\left(\phi^y\left(\cdot\right), b^{m,y}\right) \end{array}\right)\phi^y_{b^y} \tag{31}$$

We have seen that $\phi^y_{b^y} < (>)\, 0$ when $\Delta\left(b^y\right) > (<)\, 0$. Eq. (31) implies that $u_{b^y}\left(\overline{a}, b^{y,e}\right) > (<)\, 0$ and, in turn, $b^{y,e} < (>)\, b^*$ when $\Delta\left(b^y\right) > (<)\, 0$. Finally, we wish to determine the condition under which constraint (19) is slack, while constraint (20) is binding. This requires the discount factor to be sufficiently small, namely $\delta < \delta^y$ where δ^y is the solution of the following equation:

$$\frac{u\left(\underline{a}, b^{y,e}\left(\delta^y\right)\right) - u\left(\overline{a}, b^{y,e}\left(\delta^y\right)\right)}{u\left(\underline{a}, b^{m,e}\right) - u\left(\overline{a}, b^{m,e}\right)} = \frac{v^m\left(\phi^y\left(b^{y,e}\left(\delta^y\right), b^{m,e}\right), b^{m,e}\right) - \left(u^{aut} + \delta\underline{\omega}\right)}{\overline{\omega} - \underline{\omega}}$$

Clearly, when $\delta = \delta^y$, $\phi = \phi^y\left(b^{y,e}, b^{m,e}\right) = \phi^m\left(b^{m,e}\right)$ and both constraints (19) and (20) are simultaneously satisfied. ■

References

[1] Abreu, D., Milgrom, P., and D. Pearce, 1991. Information and Timing in Repeated Partnerships, *Econometrica*, 59(6), 1713-1733.

[2] Abreu, D., D. Pearce, and E. Stacchetti, 1986. Optimal Cartel Equilibria with Imperfect Monitoring, *Journal of Economic Theory*, 39(1), 251-269.

[3] Abreu, D., Pearce, D., and E. Stacchetti, 1990. Toward a Theory of Discounted Repeated Games with Imperfect Monitoring, *Econometrica*, 58 (5), 1041-1063.

[4] Acemoglu, D., and M. O. Jackson, 2015. History, Expectations, and Leadership in the Evolution of Social Norms, *Review of Economic Studies*, 82(2), 423-456.

[5] Anderlini, L., Gerardi, D. and R. Lagunoff, 2008. A "Super" Folk Theorem for Dynastic Repeated Games, *Economic Theory*, 37, 357-394.

[6] Baker, G., Gibbons, R., and K. J. Murphy, 1994. Subjective Performance Measures in Optimal Incentive Contracts, *Quarterly Journal of Economics*, 109(4), 1125-1156.

[7] Baron, D., 2001, Private Politics, Corporate Social Responsibility and Integrated Strategy, *Journal of Economics and Management Strategy*, 10, 7-45.

[8] Baron, D. P., 2010, Morally-Motivated Self-Regulation, *American Economic Review*, 100(4), 1299-1329.

[9] Benoit, J. P., and V. Krishna, 1987. Dynamic Duopoly: Prices and Quantities, *Review of Economic Studies*, 54(1), 23-36.

[10] Bernheim, B., D., and M. D. Whinston, 1998. Incomplete Contracts and Strategic Ambiguity, *American Economic Review*, 88(4), 902-932.

[11] Besley, T., and M. Ghatak, 2015. Solving Agency Problems: Intrinsic Motivation, Incentives, and Productivity, *World Development Report*.

[12] Bhaskar, V., 1998. Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems, *Review of Economic Studies*, 65(1), 135-149.

[13] Cremer, J., 1986. Cooperation in Ongoing Organizations, *Quarterly Journal of Economics*, 101(1), 33-49.

[14] Dixit, A., 1980. The Role of Investment in Entry-Deterrence, *Economic Journal*, 90(357), 95-107.

[15] Fudenberg, D., Levine, D., K., and E., Maskin, 1994. The Folk Theorem in Repeated Games with Imperfect Public Information, *Econometrica*, 62 (5), 997-1039.

[16] Fudenberg, D., and J. Tirole, 1984. The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look, *American Economic Review*, 74(2), 361-366.

[17] Fudenberg, D., and J. Tirole, 1991. Game Theory, *MIT Press*.

[18] Green, E., and R., Porter, 1984. Noncooperative Collusion under Imperfect Price Information, *Econometrica*, 52 (1), 87-100.

[19] Hammond, P., 1975. Charity: Altruism or Cooperative Egoism, in Altruism, Morality, and Economic Theory, edited by E. Phelps, *Russell Sage Foundation*.

[20] Holmstrom, B., and P. Milgrom, 1991. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design, *Journal of Law, Economics, & Organization*, 7, 24-52.

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

[22] Iannaccone, L., 1992. Sacrifice and Stigma: Reducing Free-Riding in Cults, Communes and other Collectives, *Journal of Political Economy*, 100(2), 271-291.

[23] Kandori, M., 1992. Repeated Games Played by Overlapping Generations of Players, *Review of Economic Studies*, 59(1), 81-92.

[24] Kreps, D., 1996. Corporate Culture and Economic Theory, in Firms, Organizations and Contracts: A Reader in Industrial Organization, edited by Buckley, P. J. and J. Michie, *Oxford University Press*.

[25] Lagunoff, R., and A. Matsui, 2004. Organizations and Overlapping Generations Games: Memory, Communication, and Altruism, *Review of Economic Design*, 8(4), 383-411.

[26] Levy, G., and R. Razin, 2012. Religious Beliefs, Religious Participation and Cooperation, *American Economic Journal: Microeconomics*, 4(3), 121-151.

[27] Limao, N., 2005. Trade Policy, Cross-Border Externalities and Lobbies: Do Linked Agreements Enforce More Cooperative Outcomes?, *Journal of International Economics*, 67(1), 175-199.

[28] Mailath, G., and L. Samuelson, 2006. Repeated Games and Reputations, *Oxford University Press*.

[29] Radner, R., 1986. Repeated Partnership Games with Imperfect Monitoring and No Discounting, *Review of Economic Studies*, 53 (1), 43-57.

[30] Rangel, A., 2003. Forward and Backward Intergenerational Goods: Why Is Social Security Good for the Environment?, *American Economic Review*, 93(1), 813-834.

[31] Salant, D., 1991. A Repeated Game with Finitely Lived Overlapping Generations of Players, *Games and Economic Behavior*, 3(2), 244-259.

[32] Smith, L., 1992. Folk Theorems in Overlapping Generations Games, *Games and Economic Behavior*, 4 (3), 426-449.

[33] Spagnolo, G., 1999. On Interdependent Supergames: Multimarket Contact, Concavity, and Collusion, *Journal of Economic Theory*, 89(1), 127-139.

[34] Spence, A., M., 1977. Entry, Capacity, Investment and Oligopolistic Pricing, *Bell Journal of Economics*, 8(2), 534-544.

[35] Topkis, M. D., 1998. Supermodularity and Complementarity, *Princeton University Press.*