Incentives for Collective Innovation*

Gregorio Curello†

Abstract

Agents exert effort to generate increments of random size in a flow of common benefits. I characterise the unique symmetric Markov equilibrium with benefits as the state variable, and the most efficient equilibria of the two-agent game in strategies that condition on time and the history of benefits. In the co-operative solution, any increment in benefits increases continuation payoffs. However, if enjoying benefits and augmenting them are mutually exclusive, then continuation payoffs in the Markov equilibrium may drop after an increment. Moreover, this is always true in the most efficient equilibria. As a result, free-disposal of increments contingent on their size is incentive-compatible and enhances welfare. In the context of an R&D alliance in which improvements in a collective technology raise the incentive to shift resources towards private activities, avoiding avenues of research with limited (if guaranteed) returns may enhance private profits and collective progress.

1 Introduction

In the context of innovation, adequate but imperfect solutions to a problem may do more harm than good, because they lower the incentive to find better ones. When resources are limited, such as in the social sector, there may be pressure to fund activities with guaranteed returns, however narrow and short-term they may be,

*I thank Péter Eső, John Quah, Meg Meyer, Johannes Hörner, George Georgiadis, Bruno Strulovici, Godfrey Keller, Alessio Piccolo, Ludvig Sinander, Matteo Escudé and the seminar participants at the University of Oxford for helpful comments.

†Department of Economics, University of Oxford. E-mail: gregorio.curello@economics.ox.ac.uk.

at the expense of the (costly) search for more effective interventions.[1] As far as technological innovation is concerned, late innovators may outperform the leaders, as the latter face less urgency to innovate and high switching costs.[2]

To study the incentives for collective innovation, I analyse a dynamic public good game with a stochastic production technology. At any point in time, agents exert effort to induce increments of random size in a public flow of benefits. Moreover, the opportunity cost of effort is potentially increasing in the size of benefits. I characterise the co-operative solution and the unique symmetric Markov equilibrium (with benefits as the state variable) for a wide range of payoff functions. I also construct the most efficient public-perfect and perfect equilibria of the two-player game with linear payoffs. In the former, strategies may condition on time and the past trajectory of benefits. In the latter, they may also depend on who induced each increment in benefits. Finally, I show how each equilibrium is affected by the introduction of free-disposal of increments contingent on their size.

As an example, consider an R&D alliance of firms aimed at improving a shared technology. Firms allocate their resources between public R&D and private activities. R&D induces random increments, drawn from a distribution F, in the technology level. The activities induce a private flow of profits, increasing in the level of the technology.[3]

If agents co-operate to maximise welfare, effort drops as benefits increase, but continuation payoffs always increase. Hence any innovation, however small, is welfare-enchancing. In contrast, continuation payoffs may be U-shaped in the Markov equilibrium. If so, then small innovations are harmful when benefits are small. This is because effort decreases as benefits grow, delaying future increments,

1. The non-profit organisations described by Shore, Hammond, and Celep (2013) spent years fighting some of the symptoms of poverty (earning widespread support and praise) before understanding that, by refocusing their efforts, they could address some of its root causes and have a long-lasting impact.

2. For example, mobile phone payments were introduced in 2007 in Kenya and in 2014 in the U.S. As noted by Mutiga (2014), in Kenya, '[t]he service has brought millions of people into the formal financial system [and] hobbled crime by substituting cash for pin-secured virtual accounts.'

3. There is evidence that, in the late stages of successful partnerships, progress raises the incentive to shift resources towards private activities. Powell and Giannella (2010) note that 'Once knowledge accumulates to the stage that tangible outcomes are possible, private interests may take hold and commercialise particular streams of technology that emerge from collective invention' and that 'as technological uncertainty recedes, firms develop private R&D and focus on their own specific applications. Reliance on collective invention accordingly wanes.'

and this may offset the advantage of moderately larger benefits. In contrast, a sufficiently large increment in benefits is always beneficial, even though it leads to a sharp drop in effort. If fact, continuation payoffs are high when benefits are small thanks to the prospect of one such increment, which is expected sooner the smaller the benefits, because effort is higher.

Equilibrium payoffs are U-shaped if the marginal cost of effort is low when benefits are small and high when benefits are large, and there are many agents. This is because, in this case, the first increment in benefits occurs early on and the individual cost of inducing it is low. Hence, initial payoffs are approximately the *average* of continuation payoffs after the first (random) increment. Moreover, continuation payoffs are increasing when benefits are large, since effort is low due to the high cost. This implies that continuation payoffs are not monotone, for otherwise intial payoffs would be strictly lower than the average. The U-shape arises under natural assumptions.

Since small increments in benefits are harmful, if agents could freely dispose of increments after observing their size, they would dispose of the small ones, and this would improve efficiency. Moreover, if free-disposal is not allowed, improving the distribution F may be detrimental as it may delay the growth of benefits in the medium run, and this may offset its short- and long-run advantages. Specifically, FOSD-shifts of F that reduce the risk of negligible increments in favour of moderate ones reduce initial expected discounted payoffs.

In the context of the R&D alliance, if the profits of private activities are sensitive to the efficiency of the shared technology, then small technological improvements may be harmful in the early stages of the alliance. In particular, if a number of separate alliances work on the improvement of the same technology, a network of late innovators is likely to leapfrog into a leading position (that is, reach larger benefits sooner) if the former leaders produced innovations of moderate value. Moreover, firms may choose to disregard an avenue of research once it has become clear that its product will be moderate (that is, to dispose of a small increment in benefits), in order to preserve their parterns' incentives to R&D. Finally, if it is commonly agreed that each technological improvement will be immediately monetised (that is, if free-disposal never occurs), then improvements in the productivity of R&D which reduce the risk of unproductive research but do not increase the odds of obtaining radical breakthroughs may slow down progress.

In the two-player game, time and the past trajectory of benefits cannot incentivise effort more than the current size of benefits. Indeed, in the most efficient public-perfect equilibrium, aggregate effort is pinned down by current benefits. However, this equilibrium is not Markov, as the past trajectory of benefits is a useful correlation device. Only one agent exerts effort after benefits exceed a certain cutoff and, before this occurs, each agent has a one-in-two chance of being the one to do so. Efficiency can be improved further if effort may condition on who induced each past increment. In the most efficient perfect equilibrium of the game beginning at intermediate sizes of benefits, the agent who induces the first increment free-rides for the remainder of the game.

In both equilibria, aggregate payoffs are likely to drop after the last increment in benefits unless it is significant. Thus, innovations of itermediate size are harmful in the early stages of the alliance, and small innovations are detrimental in the late stages. Moreover, agents' incentives to dispose of increments are initially aligned, but become misaligned once an agent stops exerting effort. Despite this, allowing free-disposal is beneficial once again. In particular, agents dispose of an increment in benefits if and only if doing so is advantageous to both.

This dynamic game is mathematically novel as the size of jumps of the state has an arbitrary distribution. Agents' value functions solve integral equations and cannot be expressed analytically. Moreover, I use original dynamic programming techniques to compute the most efficient equilibria within families that are too rich to characterise in full.

The rest of the paper is organised as follows. Section 2 contains a review of the related literature. The model is laid out in Section 3. In Section 4, I present the co-operative solution and the unique symmetric Markov equilibrium. In Section 5, I characterise the most efficient public-perfect equilibria of the two-player game, as well as the most efficient perfect equilibria for sufficiently large initial benefits. In Section 6, free-disposal and comparative statics with respect to F are discussed.

2 Literature review

Our model builds on the literature of dynamic games of private provision of public goods in which the good is accumulated over time. Fershtman and Nitzan (1991) study one such model and exhibit a symmetric MPE in which the long-run stock of the good is inefficiently low due to the inter-temporal free-riding effect. This occurs in *all* MPE in our model. Admati and Perry (1991) derive a similar result for a game in which players contribute in sequence to a public project. Even though players act simultaneously in our model, they take turns at exerting effort before ceasing completely in the most efficient MPE. Marx and Matthews (2000) consider a similar model and analyse Markov equilibria as well as SPNE, which allow players' actions to be conditioned on opponents' past actions. They show that an approximately efficient SPNE arises as the period length tends to zero. Although I do not prove it, this is also the case in our model. Battaglini, Nunnari, and Palfrey (2014) constrain the investments in the public good to be irreversible and analyse symmetric and continuous MPE. They show that the (unique) steady-state of the best equilibria converges to the efficient steady-state as time discounting vanishes, and the set of all steady-states shrinks to a unique point as depreciation vanishes. Investments in our model are also irreversible and these results hold for *all* MPE. All of the models above are deterministic. Wang and Ewald (2010) add a diffusion term to the stock of the public good in the model of Fershtman and Nitzan (1991). In the specification that is closest to our model, the variance of the diffusion is increasing in the agents' aggregate effort. The authors show that higher variance exhacerbates the free-riding effect. In contrast, I show that improvemets in productivity may exhacerbate free-riding (even when, for any fixed strategy profile, these improvements reduce the variance of the process describing the stock of the public good).

Our model is close to the literature on games of social experimentation, which largely relies on bandit models. In these games, players allocate their time between pulling a 'safe' arm, which generates a flow of payoffs from a known stochastic process, and a 'risky' arm, generating payoffs from an uncertain process. Pulling the risky arm is informative about its returns: for this reason, doing so is labelled 'experimenting'. Payoffs are fully observable, hence experimentation produces information spillovers. In our model, agents dynamically choose between a passive costless stance and a costly 'experiment' with random returns. However, the stochastic process generating these returns is fully known, hence information has no value. The strategic interaction is based on the fact that, unlike in bandit models, *payoffs* are public.

In the exponential bandit model of Keller, Rady, and Cripps (2005), the risky arm is either better than the safe one, or it generates no payoffs. Hence, in the absence of payoffs, the belief about the quality of the risky arm is a monotonic function of time. Similarly, the flow of benefits increases with time in our model. This allows to obtain MPE by backward induction. The authors show the existence of a uniformly most efficient MPE when agents exert either full or no effort, and may switch between the two only finitely many times. I show that our model admits a unformly most efficient MPE when players' strategies are unrestricted. This MPE is qualitatively similar to an MPE presented by Keller and Rady (2010). They study a model of Poisson bandits in which the state is non-monotonic, and a uniformly most efficient MPE is unlikely to exist. Their model is more closely related to the seminal model of Bolton and Harris (1999), as all MPE feature the encouragement effect. In contrast, this effect never arises in our model.

Several papers build on these bandit models to study incentives for experimentation. Bonatti and Hörner (2011) study a game similar to that of Keller, Rady, and Cripps (2005), but in which payoffs are shared. They show that imposing a 'deadline' after which no experimentation can occur increases efficiency, whether players play the (unique) symmetric MPE or the symmetric SPNE arising when effort is unobserved. Halac, Kartik, and Liu (2016) use the exponential bandit model to compare two natural prize and information-sharing schemes in contests for experimentation. Our comparative statics results suggest that seemingly beneficial incentives may have adverse effects in our model. I plan to pursue this question by allowing the agents to choose between distinct production technologies at any point in time. Garfagnini and Strulovici (2016) study a richel model of social experimentation that differentiates between 'radical' and 'incremental' innovations, where the former expand the set of experiments available to the agents. FOSD- and MLR-shifts of the distribution F in our model (which have opposite welfare effects, as noted in the introduction), can be thought of as arising from 'incremental' and 'radical' innovations, respectively, in related areas.

3 Model

In this section, I lay out the model in its general form as well as the special example of the R&D alliance.

Framework. Time is continuous and indexed by $t \in [0, \infty)$. Each of N agents, indexed by $i \leq N$, exerts costly effort to increase a flow of common benefits. Let $x_t > 0$ denote (the intensity of the flow of) benefits at time $t \geq 0$. In any interval of time $[t, t + \mathrm{d}t)$, Agent i exerts effort $a_t^i \in [0, 1]$ and receives payoff $u(a_t^i, x_t)$, discounted at rate $r > 0$. Assume that u is continuous, $u(\cdot, x)$ is strictly decreasing and concave, $u(a, \cdot)$ is strictly increasing and concave, $u_1(a, \cdot)$ is increasing, and $u_2(\cdot, x)$ is decreasing.

Benefits x_t take some initial value $x_0 \geq 0$ and, for any $t > 0$, are determined as follows. In any interval of time $[t, \mathrm{d}t)$, each Agent i produces an increment in x_t with probability $a_t^i \mathrm{d}t$. Each increment has random size s, drawn from a distribution F with support within $[0, \infty)$ and mean $\mu < \infty$. The production and the size of increments are independent across agents, across time, and from one another.[4]

Baseline example: R&D partnership. In this example, agents are firms and x_t represents the current performance of the technology that they share. Firm i devotes a fraction a_t^i of its resources to R&D aimed at improving the technology, and the rest to private activities. Assume that the flow benefit to the firm from doing so is given by $x_t(1 - a_t^i)\mathrm{d}t$, so that $u(a, x) = x(1 - a)$.

4. Equivalently, with probability $(\sum_{i=1}^N a_t^i)\mathrm{d}t$, $x_{t+\mathrm{d}t} = x_t + s$ where s is a random draw from F; if this does not occur, then $x_{t+\mathrm{d}t} = x_t$. Note that replacing $(\sum_{i=1}^N a_t^i)\mathrm{d}t$ with $\lambda(\sum_{i=1}^N a_t^i)\mathrm{d}t$ for some $\lambda > 0$ and resizing the domain of a_t^i to $[0, \alpha]$ for some $\alpha > 0$ is strategically equivalent to replacing r with $r/(\lambda\alpha)$ and $u(\cdot, x)$ with $u(\cdot/\alpha, x)$.

4 Welfare benchmark and Markov equilibrium

The model laid out in the previous section is a dynamic game with state variable x_t. In this section, I compute the welfare benchmark and the (unique) symmetric Markov equilibrium of the game. Among other things, I show that continuation payoffs in the symmetric equilibrium need not be increasing in x_t.

Welfare is the average of expected discounted payoffs across agents in the game with initial size of benefits x_0, denoted by $w(x_0)$. A *Markov strategy* is a measurable function $\sigma^i : [0,\infty) \to [0,1]$ such that $a_t^i = \sigma^i(x_t)$ for all $t \geq 0$. Since this is a dynamic game with state x_t, welfare is maximised by a profile of Markov strategies. The following result shows that, in order to maximise welfare, agents should exert less effort as benefits grow. It is proved in the Appendix.

Theorem 1 (Co-operative solution)**.** *Welfare is maximised if agents play a symmetric, continuous, decreasing strategy profile σ^*. Maximal welfare w^* is increasing and continuous in initial beneftis x_0.*

The decreasing efficient effort schedule is standard in dynamic public good games, and is generally due to strictly concave payoffs. In this model, not only are payoffs (weakly) concave, but the opportunity cost of effort is increasing (i.e. $u_1(a,\cdot)$ is decreasing). The latter force adds to the former and does not alter the qualitative features of the co-operative solution. However, as we shall see, it plays an important role in the symmetric Markov equilibrium.

If σ^* reaches 0 at some $\hat{x} \in (0,\infty)$ then

$$N \int_0^\infty u(0, \hat{x}+s)\mathrm{d}F(s) - u(0,\hat{x}) = r u_1(0,\hat{x}). \tag{1}$$

The left-hand side of (1) is the social value of an increment in benefits when their size is $\hat{x}$; the right-hand side is the marginal cost of producing an increment. In the baseline example, σ^* reaches 0 at $\hat{x} = N\mu/r$. Moreover, since the cost of effort is linear, σ^* exhibits a bang-bang feature: $\sigma^*(x) = N$ for $x < \hat{x}$, and $\sigma^*(x) = 0$ for $x > \hat{x}$.

We turn to Markov equilibria. A profile σ of Markov strategies forms a *Markov Perfect Equilibrium* (MPE) if each strategy is a best-response to the others in the game starting at state x_0, for any $x_0 \geq 0$. The following result characterises the unique symmetric MPE of the game. It is proved in the Appendix.

Theorem 2 (Symmetric MPE)**.** *There exists a unique symmetric MPE σ_M. Effort is continuous and decreasing in benefits x_t. Welfare w_M is continuous in x_0. Moreover, if $u_{12} \equiv 0$ then w_M is increasing. However, if $u_1(\cdot, 0) \equiv 0$ then, for sufficiently large N, w_M is non-monotone.*

If $u_{12} \equiv 0$ (e.g. if $u(a, x) = x - a$), then the opportunity cost of effort does not increase with benefits. In particular, there is no trade-off between enjoying the public good and producing it. In this case, even though less effort is exerted as benefits increase (σ_{M} is decreasing), so that further increments are delayed, the extra benefits overcome this delay, raising continuation payoffs. This is a standard feature of symmetric equilibria of dynamic public good games.

If $u_1(\cdot, 0) \equiv 0$, then the cost of effort becomes arbitrarily small as benefits shrink to 0. Essentially, this means that if benefits are low then the incentive to produce is very high.[5] This is the case in the baseline example, since $u_1(a, x) = -x$. To see why welfare is not monotone in this case, note that if interior effort is exerted when the size of benefits is $x > 0$ (i.e. $0 < \sigma(x) < 1$) then

$$\int_0^\infty w(x + s)\mathrm{d}F(s) - w(x) = ru_1(\sigma(x), x). \tag{2}$$

This expression is similar to (1); the left-hand side is the expected (individual) value of an increment in benefits given that σ will be played after the increment, and the right-hand side is the marginal cost of producing an increment. Letting x tend to 0 we obtain

$$w(0) \approx \int_0^\infty w(s)\mathrm{d}F(s). \tag{3}$$

Essentially, this means that, since the cost of effort is negligible when benefits are very low, a lot of effort is exerted so that benefits increase almost immediately. If w were monotone, then (3) would hold with '$\leq$' and equality would hold if and only if w is constant. Since w is not constant, it follows that it is not monotone.

5. Notice that this condition is strong, however it guarantees non-monotonicity regardless of the values of the rest of the parameters. Moreover, by continuity, increasing $u_1(\cdot, 0)$ by a sufficiently small amount does not restore monotonicity as long as the rest of the parameters remain fixed.

The condition that N is sufficiently large is required because (2) holds with '$\geq$' if $\sigma(x) = 1$. If $u_1(\cdot, 0) \equiv 0$ then, for any given N, $\sigma(x) = 1$ if and only if $x \leq \epsilon$ for some $\epsilon > 0$. However, as N grows large, ϵ shrinks to 0.

Despite the generality of Theorem 2, the mechanism producing the non-monotonicity is the same across all specifications of the model. To see why this is the case, note that for any two benefit sizes $x' > x > 0$ such that $w(x) > w(x')$, some effort must be exerted at x, so that (2) holds with '$\geq$' (and with '$=$' if interior effort is exerted). Then,

$$\int_{x'-x}^{\infty} w(x+s) - w(x)\mathrm{d}F(s) \geq ru_1(\sigma(x), x) - \int_{0}^{x'-x} w(x+s) - w(x)\mathrm{d}F(s).$$

Each integrand is the value from an increment of size s in benefits their size is x. On the left-hand side, we integrate over all sizes s increasing benefits above x', and we intergrate over the rest on the right-hand side. As part of the proof of Theorem 2, I show that w is decreasing and then increasing (it is convex and then concave, and it is convex in the baseline example). Then, the integral on the right-hand side is negative. This means that, at x, the benefit from the prospect of overtaking x' with the next increment exceeds the cost of inducing it. Hence, not only is there a strictly positive probability of leapfrogging any larger size of benefits inducing lower continuation payoffs – the likelihood and the benefit of this event are sufficiently large to justify the investment in production.

The mechanism inducing the non-monotonicity suggests that discrete increments in benefits are crucial to obtain the result, as they allow the possibility of leapfrogging the worst states to reach more efficient ones. In fact, the stochastic size of jumps is also crucial. If F puts all mass on a single value $\mu > 0$ then, although w remains non-monotone (Theorem 2 still applies), the trajectory $(w(x_t))_{t\geq 0}$ is monotone with probability 1, regardless of initial benefits x_0.

The fact that welfare is increasing in the co-operative solution highlights the importance of strategic incentives. Equilibrium welfare may drop as benefits increase because effort drops inefficiently fast. This kind of inefficiency is standard in dynamic public good games, and it is caused by *inter-temporal free-riding*: since σ

is decreasing, agents are reluctant to exert effort as this causes their opponents to exert less effort in the future.

I end this section with a discussion of the encouragement effect. If effort in the symmetric MPE ceases at some benefit size $\hat{y} \in (0, \infty)$, then $\hat{y}$ solves

$$\int_0^\infty u(0, \hat{y} + s)\mathrm{d}F(s) - u(0, \hat{y}) = ru_1(0, \hat{y}). \tag{4}$$

In the baseline model, $\hat{y} = \mu/r$. A comparison between (4) and (1) shows that effort ceases exactly when it stops in the single-agent problem; that is, there is no encouragement effect. As we shall see later, this does not hinge on the restriction to symmetric equilibria, or to Markov strategies. As long as effort is unobservable, no perfect equilibrium features the encouragement effect. This contrasts the findings of Keller, Rady, and Cripps (2005): in their model, the encouragement effect does not arise in the symmetric MPE, but it occurs in another, asymmetric, MPE. The difference between the equilibria is that, in the latter, agents take turns at exerting effort and switch infinitely many times as the state approaches the stopping threshold. This cannot occur in our model since agents have no control over the size of the increments in benefits.[6]

5 Public- and Subgame-Perfect Equilibria

In this section, I characterise the most efficient public- and subgame-perfect equilibria of the two-player game with linear cost of effort. I show that, in the public-perfect equilibrium, efficiency can be achieved with dynamic consistency. I also show that both equilibria always exhibit a non-monotonicity in payoffs, as the one that may arise in the symmetric MPE.

Let us begin with Public-Perfect Equilibria (PPE). In PPE, agents may condition their effort on time and on the past trajectory of benefits. PPE allow for a much richer set of behaviours than MPE, as the latter compel effort to depend on current benefits alone. In particular, PPE may generate stronger incentives to exert effort by having agents play a less efficient equilibrium if benefits do not increase for

6. Moreover, I conjecture that this is not in itself the reason preventing the encouragement effect.

a sufficiently long period of time, or if they have increased too slowly in the past. As we will see, this device is not useful: the most efficient PPE does not feature it.

Formally, a *Public-Perfect Equilibrium (PPE)* is a perfect equilibrium in strategies such that a_t^i may be written as a function of the *public history* – the set $\{x_s : s \in [0,t)\}$ if $t > 0$, or the value of x_0 if $t = 0$.[7] In this context, *welfare* (the average of players' expected discounted payoffs) is a function of the public history, not merely of current benefits. The next result characterises the most efficient PPE of the game. It is proved in the Appendix.

Theorem 3 (PPE)**.** *If $N = 2$, $u(\cdot, x)$ is linear, and F has no atoms, there exists a PPE σ_P inducing highest welfare among PPE after any history. There exists $\hat{z} < \hat{y}$ such that, in any such PPE*

1. *Aggregate effort $\bar{e}(x_t)$ is highest among PPE and is pinned down by current benefits x_t: it is decreasing and drops to 0 at $x_t = \hat{y}$.*
2. *Welfare $\bar{w}(x_t)$ is pinned down by x_t: it is continuous and increasing on $[\hat{z}, \hat{y}]$ and on $(\hat{y}, \infty)$, but jumps down at $\hat{y}$.*

Moreover, σ_P can be chosen so that agents' effort and continuation payoffs are symmetric whenever $x_t \leq \hat{z}$. Finally, aggregate effort and welfare in σ_P can be approximated by MPE; that is, there exists a sequence of MPE $(\sigma_n)_{n=1}^{\infty}$ such that $\lim_n \sup_x |\sum_{i=1}^{2} \sigma_n^i(x) - \bar{e}(x)| = \lim_n \sup_x |w_{\sigma_n}(x) - \bar{w}(x)| = 0.$

The existence of an equilibrium that is the most efficient one after any history is by no means guaranteed. In fact, existence often fails even if initial benefits x_0 are fixed and attention is restricted to histories on the equilibrium path. In our model, for small enough x_0, no such equilibrium exists among subgame-perfect equilibria, and among MPE.[8] This is because one can provide strong incentives to act efficiently by having agents play an inefficient continuation equilibrium after histories that are more likely to be reached after inefficient play. Yet, this notion of efficiency is necessary to be able to maximise aggregate payoffs in a time-consistent

7. The set of such strategies is simple as x_t is right-continuous, increasing, and takes finitely many values on any bounded interval.

8. As Theorem 4 shows, a Markov equilibrium is among the most efficient SPNE if x_0 is sufficiently close to $\hat{y}$, or above it.

manner. If agents may co-ordinate on a different equilibrium in each period, then the notion of efficiency purely based on expected discounted payoffs at the beginning of the game is inappropriate.

Examples of equilibria that are the most efficient ones after every history are rare in the literature on dynamic games. One reason is that their existence is ruled out if the trajectory of the state is not monotone.[9] An example is given in Keller, Rady, and Cripps (2005), in which the state is essentially monotone. They characterise the MPE of the two-player game in pure strategies involving finitely many switches between the safe and the risky arm, and show that they can be ranked according to use of the risky arm at all states, and equilibria involving more experimentation at all states are more efficient at all states. This immediately implies the existence of an equilibrium that is most efficient at all states.

The argument in the proof of Theorem 3 is different. I do not characterise the full set of PPE, as it is very rich.[10] Instead, I show that the highest achievable welfare in PPE, as a function of the initial state x_0, ($w(x_0)$, say) lies below the solution of a particular Bellman equation. I then solve the equation and show that the solution corresponds to the welfare function of a particular PPE σ_{P}. The main step in the construction of the equation is to note that whenever Agent i exerts effort in some PPE σ, her continuation payoff v_i satisfies

$$\int_0^\infty \bar{v}_i(x_t + s)\mathrm{d}F(s) - v_i \geq rx_t$$

where $\bar{v}_i(x_t + s)$ is her continuation payoff after an increment of size s occurs at time t. If agents exert effort simultaneously then this holds for $i = 1, 2$, so that

$$\frac{v_1 + v_2}{2} - rx_t \leq \int_0^\infty \frac{\bar{v}_1(x_t + s) + \bar{v}_2(x_t + s)}{2}\mathrm{d}F(s) \leq \int_0^\infty w(x_t + s)\mathrm{d}F(s). \quad (5)$$

If there exist a sequence of σ's inducing welfare that approximate $w(x_t)$, then (5) implies that

$$w(x_t) - rx_t \leq \int_0^\infty w(x_t + s)\mathrm{d}F(s).$$

9. See the discussion in Keller and Rady (2010).
10. In particular, they cannot be ranked according to effort or welfare after every history.

If no such sequence exists, then since $N = 2$, $w(x_t)$ must be approximated by PPE in which only one agent exerts effort when benefits have size x_t. The fact that the cost of effort is linear implies that this agent must exert maximal effort. This induces another bound on $w(x_t)$. Combining the bounds produces the Bellman equation.

Even though PPE allow behaviour to depend on time and on the whole past trajectory of benefits, aggregate effort $\bar{e}$ and welfare $\bar{w}$ in the most efficient PPE σ_{P} merely depend on x_t. In particular, time and the past trajectory of x_t do not provide useful direct incentives for effort. Yet, σ_{P} is not a MPE. That is because the scope for co-ordination is larger in PPE than in MPE, and this matters for welfare. In particular, in σ_{P}, only one agent exerts (maximal) effort when the size of benefits lies within $[\hat{z}, \hat{y}]$, and her identity is not a function of current benefits. Rather, when benefits are smaller than $\bar{z}$, each agent faces a 50-50 chance of having to exert effort in the interval $[\hat{z}, \hat{y}]$ once benefits enter it.[11] This device generates symmetric continuation payoffs at benefits smaller than $\hat{z}$, and this is crucial to maximise effort, and therefore welfare. Put differently, the most efficient PPE is also the most equitable PPE from an ex-ante perspective.[12]

Welfare $\bar{w}$ is not monotone, but for a different reason than the one breaking monotonicity in the symmetric MPE σ_{M}. Welfare in σ_{M} is not monotone if the cost of effort is negligible when benefits are small (i.e. $u_1(\cdot, 0) = 0$), whereas $\bar{w}$ is not monotone if the cost of effort is linear. In particular, if $u(a, x) = x - a$, then welfare is monotone in σ_{M}, but it is not monotone in σ_{P}. Note that, as shown in Figure 1, σ_{P} is similar to σ_{M} below $\hat{z}$, and it equals the solution to the single-agent problem above $\hat{z}$.[13] The latter is more efficient than any symmetric equilibrium in a neighbourhood of $\hat{y}$ if the cost of effort is not too convex (in particular, if it is linear). Welfare drops at $\hat{y}$ because the continuation payoffs of the agent exerting effort in $[\hat{z}, \hat{y}]$ are continuous in a neighbourhood of $\hat{y}$ and, since she ceases to exert

11. Markov strategies do not allow this because the probability of benefits increasing to some $x \in [\hat{z}, \hat{y}]$ varies with their current size. Moreover, a fair lottery can only be induced if F is atomless.

12. The same is true in Keller, Rady, and Cripps (2005).

13. This equilibrium is qualitatively similar to the Markov equilibrium described in Section 6 of Keller and Rady (2010). Welfare is monotone in their equiibrium because, in games of experimentation, welfare and the incentive to exert effort are increasing in the state. In contrast, in classical public good games as well as in this model, the incentive to exert effort is decreasing in the state.

effort abruptly, the continuation payoffs of her opponent drop.

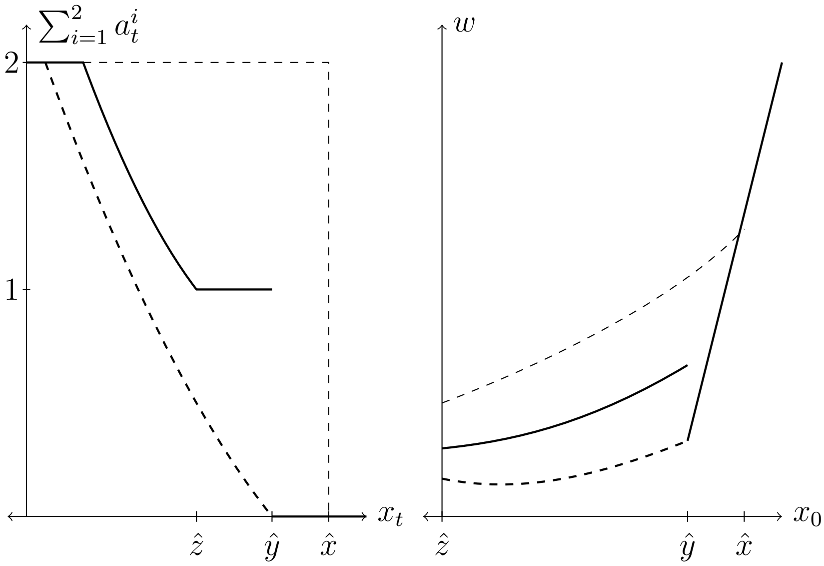

Figure 1: Aggregate effort (left) and welfare (right) in the symmetric MPE (thick dashed line), the most efficient PPE (thick line), and the welfare benchmark (thin dashed line), as functions of benefits.

We turn to Subgame-Perfect Nash Equilibria (SPNE), and maintain the assumption that agents' effort is not obervable to their opponents.[14] In SPNE, agents may not only condition their effort on time and the past trajectory of benefits, but also on who generated each increment. I show that this information can be used to induce stronger incentives for effort, improving efficiency over any PPE. In the most efficient SPNE with intermediate initial benefits, the first agent who induces an increment is rewarded with the right to free-ride (i.e. exert no effort) for the rest of the game, while her opponent essentially plays the game alone.

In this setting, efficiency cannot be maximised in a dynamically consistent manner. That is, for some values of x_0, there does not exist a SPNE that is the most efficient one after every history on the equilibrium path. Indeed, if the above equilibrium is played, it is clear that after a sufficiently small increment in benefits it would be more efficient to have another round of competition, and only allow free-riding once benefits increase again.

Formally, a *Subgame-Perfect Nash Equilibrium* is a perfect equilibrium in strate-

14. One can show that, if effort is observable, there exists an efficient SPNE; that is, one inducing maximal welfare (c.f. Theorem 1).

gies such that a_t^i may be written as a function of the *history* of play; that is, a pair $h_t = \left(h_t^p, (i_k)_{k=1}^K\right)$ where h_t^p is a public history in which x_s has K discontinuities, and $i_k \in \{1, \ldots, N\}$ for all $k \leq K$. Each discontinuity of x_s marks the production of the public good, and i_k tracks the identity of the producer. In this context, I focus on welfare (average of agents' expected discounted payoffs) at the beginning of the game, and disregard welfare at later stages. The following result characterises the most efficient SPNE assuming that initial benefits are sufficiently high. It is proved in the Appendix.

Theorem 4 (Theorem)**.** *There exists $\tilde{z}' < \hat{y}$ such that, if $N = 2$, $u(\cdot, x)$ is linear, and $x_0 \geq \tilde{z}'$, then the following SPNE σ_{N} induces highest welfare among SPNE:*

- *If $x_0 > \hat{y}$, no effort is exerted.*
- *For some $\tilde{z} \in (\tilde{z}', \hat{y})$, if $x_0 \in (\tilde{z}, \hat{y}]$, then one player exerts (full) effort until x_t exceeds $\hat{y}$, and the other exerts no effort.*
- *If $x_0 \in [\tilde{z}', \tilde{z}]$, then players exert constant symmetric effort until x_t increases. After the increase, the player who generated it stops, whereas her opponent exerts (full) effort until x_t exceeds $\hat{y}$.*

6 Free-disposal and comparative statics

In this section, I study the impact of allowing agents to dispose of increments in x_t after observing their size, and of altering the size distribution F. Among other things, I show that free-disposal restores the monotonicity of payoffs in the symmetric MPE σ_{M}, and that forcing the agents to dispose of sufficiently small increments enhances ex-ante welfare in the most efficient PPE σ_{P} and SPNE σ_{N}.

Free-disposal. Formally, free-disposal is defined as follows. I assumed in Section 3 that, in any interval of time $[t, \mathrm{d}t)$, for any $i \leq N$, Agent i produces an increment with probability $a_t^i \mathrm{d}t$, and the increment has random size s drawn from F. In the *game with free-disposal*, in any interval of time $[t, t + \mathrm{d}t)$, for any $i \leq N$, a value s is drawn from F with probability $a_t^i \mathrm{d}t$. If this occurs, Agent i observes s and decides whether to increase x_t to $x_t + s$.[15]

15. More formally, agents choose, for each time t, a measurable set $S_t^i \subseteq [0, \infty)$ such that, if a value s

Welfare benchmark. The co-operative solution is unaffected by the introduction of free-disposal, as the latter is never efficient. That is because any increment in x_t is strictly beneficial, regardless of its size. Moreover, the larger the size of the increment, the larger its benefit. Hence any FOSD-shift of F increases welfare in the co-operative solution. This also increases the expected benefit of any increment, so that effort rises too.[16]

Symmetric MPE. If free-disposal is allowed, then agents dispose of an increment if and only if it would reduce their continuation payoffs. I showed in Section 4 that, in the symmetric MPE of the game without free-disposal, any increment is beneficial if $u_{12} \equiv 0$, whereas small increments are detrimental when x_t is small if $u_1(\cdot, 0) \equiv 0$. As a consequence, free-disposal does not occur in the former case, but occurs in the latter. Moreover, when free-disposal is allowed, increments that are not disposed of increase continuation payoffs. These results are summarised in the next proposition.[17]

Proposition 1 (Free-disposal)**.** *Theorem 2 continues to hold in the game with free-disposal. If $u_{12} \equiv 0$, free-disposal never occurs and σ_M is unaffected. If $u_1(\cdot, 0) \equiv 0$, free-disposal increases welfare in σ_M, and strictly so for sufficiently small x_0. Moreover, in the game with free-disposal, $w(x_t)$ is increasing in t.*

Next, I show that, if free-disposal is not allowed, then improvements in F have ambiguous effects in the symmetric MPE.[18] In particular, MLR-shifts of F are always beneficial but, if welfare is not monotone, then some FOSD-shifts of F are detrimental. The harmful FOSD-shifts improve F on the lowest part of its support. That is, they reduce the risk that x_t increases by very small amounts, and increase the likelihood that it increases by intermediate amounts; however, they do not improve the odds of generating large increments in x_t.

It is clear that, if equilibrium welfare is not monotone, then *short-term* FOSD-improvements may be detrimental. Namely, if weflare decreases in a neighbourhood of the initial benefits, then improving the distribution of the *first* increment over

is drawn from F at time t, then x_t is increased to $x_t + s$ if and only if $s \in S_t^i$.

16. The proofs of these results are omitted.

17. The game with free-disposal admits a unique symmetric MPE. As before, effort σ_{M} is continuous and decreasing in x_t. Proofs of these results are omitted.

18. If free-disposal *is* allowed, then any FOSD-shift of F increases equilibrium welfare.

this neighbourhood decreases initial weflare.[19] However, a permanent improvement of F is beneficial in the long term. That's because, as we have seen, equililbrium welfare is increasing in benefits if the latter are sufficiently large. We will see that the short-term negative effects outweight the long-term positive ones: some FOSD-shifts of F may decrease expected discounted payoffs at the beginning of the game.

Given F and $\epsilon > 0$, let

$$F_\epsilon(x) = \begin{cases} 0 & \text{if } x < \epsilon \\ F(x) & \text{otherwise.} \end{cases}$$

Shifting F to F_ϵ pushes all the mass that F puts on the interval $[0, \epsilon)$ onto $\{\epsilon\}$. This means that the shift guarantees increments of size at least ϵ, but has no effect on the size of an increment conditional on it exceeding ϵ. The next result describes how improving F affects the symmetric MPE σ_{M}. It is proved in the Appendix.

Proposition 2 (Improvements in F)**.** *In the game without free-disposal, any MLR-shift of F increases welfare in σ_M. However, if $u_1(\cdot, 0) = 0$, then for sufficiently small $\epsilon > 0$, shifting F to F_ϵ decreases welfare in σ_M if N is sufficiently large and x_0 is sufficiently small.*

Most efficient PPE & SPNE. In asymmetric equilibria, the introduction of free-disposal may have adverse effects as an agent who disposes of an increment may reduce the continuation payoffs of her opponent. Despite this, allowing free-disposal strictly increases welfare in the most efficient PPE σ_{P} and in the most efficient SPNE σ_{N}.[20] This is because, despite being asymmetric, these equilibria have a particular feature. Namely, as long as both agents exert effort, their incentives are aligned. Hence, if one finds it profitable to dispose of an increment, the other will benefit too. Thus, free-disposal is beneficial at early stages of the game. Moreover, once an agent stops exerting effort, the other will never find it profitable to dispose of any increment. Hence, free-disposal has no effect in the late stages of the game. These

19. A formal proof of this statement involves an extended model in which F may depend on x_t, so that F is improved if $x_t = x_0$, but remains fixed for any $x_t > x_0$.

20. In this section, I maintain the assumptions under which the equilibria σ_{P} and σ_{N} were computed. Namely, I assume that $N = 2$, F has no atoms, and $u(a, x) = x(1 - a)$. Moreover, when referring to σ_{N}, I implicitly assume that $x_0 \geq \tilde{z}'$ (c.f. Theorem 4).

results are summarised in the following proposition. It is proved in the Appendix.

Proposition 3 (Free-disposal)**.** *Theorems 3 and 4 continue to hold in the game with free-disposal. Free-disposal increases weflare in* σ_P *(resp.* σ_N*), and strictly so for* $x_0 < \hat{z}$ *(resp.* $x_0 < \tilde{z}$*). Moreover, free disposal only occurs when both agents are exerting effort.*

References

Admati, A., and M. Perry. 1991. "Joint projects without commitment." *The Review of Economic Studies* 58 (2): 259–276.

Battaglini, M, S Nunnari, and T R Palfrey. 2014. "Dynamic free riding with irreversible investments." *American Economic Review* 104 (9): 2858–71.

Bolton, P., and C. Harris. 1999. "Strategic experimentation." *Econometrica* 67 (2): 349–374.

Bonatti, Alessandro, and Johannes Hörner. 2011. "Collaborating." *American Economic Review* 101 (2): 632–63.

Fershtman, C., and S. Nitzan. 1991. "Dynamic voluntary provision of public goods." *European Economic Review* 35 (5): 1057–1067.

Garfagnini, Umberto, and Bruno Strulovici. 2016. "Social experimentation with interdependent and expanding technologies." *Review of Economic Studies* 83 (4): 1579–1613.

Halac, Marina, Navin Kartik, and Qingmin Liu. 2016. "Optimal contracts for experimentation." *The Review of Economic Studies* 83 (3): 1040–1091.

Keller, G., and S. Rady. 2010. "Strategic experimentation with Poisson bandits." *Theoretical Economics* 5 (2): 275–311.

Keller, G., S. Rady, and M. Cripps. 2005. "Strategic experimentation with exponential bandits." *Econometrica* 73 (1): 39–68.

Marx, L., and S. Matthews. 2000. "Dynamic voluntary contribution to a public project." *The Review of Economic Studies* 67 (2): 327–358.

Mutiga, M. 2014. "Kenya's Banking Revolution Lights a Fire." *The New York Times.* `https://www.nytimes.com/2014/01/21/opinion/kenyas-banking-revolution-lights-a-fire.html`.

Powell, Walter W, and Eric Giannella. 2010. "Collective invention and inventor networks." In *Handbook of the Economics of Innovation,* 1:575–605. Elsevier.

Shore, B, D Hammond, and A Celep. 2013. "When Good Is Not Good Enough." *Stanford Social Innovation Review.* `https://ssir.org/articles/entry/when_good_is_not_good_enough`.

Wang, Wen-Kai, and Christian-Oliver Ewald. 2010. "Dynamic voluntary prodivion of public goods with uncertainty: a stochastic differential game model." *Decisions in Economics and Finance* 33 (2).

Appendix

Given $w : [0, \infty) \to \mathbb{R}$, $x \geq 0$ and $a \in [0, 2]$, define the functional

$$Tw(x, a) = \frac{x + a(\mathbb{E}_x w - c_x/2)}{a + r}. \tag{6}$$

Lemma 1. *Given $x_0 > 0$ and $w_0 : [x_0, \infty) \to \mathbb{R}$ such that $w_0(x_0) = Tw_0(x_0, 2)$, there exists a unique $w : [0, \infty) \to \mathbb{R}$ that is continuous on $[0, x_0]$ and such that, for all $x \geq 0$*

$$w(x) = \begin{cases} w_0(x) & \text{if } x > x_0 \\ Tw(x, 2) & \text{otherwise.} \end{cases}$$

Morever, if $\delta(x_0) \geq 0$ *and* $w_0(x) \geq x/r$ *for all* $x \geq x_0$*, then* $w(x) \geq x/r$ *for all* $x \geq 0$*. If* w_0 *is increasing, then* w *is increasing. If* w_0 *satisfies both of the above properties and, furthermore, the rate of change of* w_0 *is bounded by* $1/r$*, then so is the rate of change of* w*.*

Proof. The result is proved through repeated applications of the contraction mapping theorem. To prove the first part, let $I = [0, x_0]$ and let C be the set of continuous functions $w : I \to \mathbb{R}$. Note that C is a complete metric space with the supremum norm. Given $w \in C$ and $x \in I$, let

$$T_0 w(x) = \frac{x + 2\mathbb{E}_x \left[\mathbb{I}_{s \leq x_0} w(s) + \mathbb{I}_{s > x_0} w_0(s) \right] - c_x}{2 + r} \tag{7}$$

Since $\mathbb{E}_x$ and w are continuous on I, then so is $T_0 w$. Hence T maps C to itself. Moreover, for any $v, w \in C$,

$$\sup_{x \in I} |T_0 v(x) - T_0 w(x)| = \frac{2}{2+r} \sup_{x \in I} |E_x \mathbb{I}_{s \leq x_0} [v(s) - w(s)]| \leq \frac{2}{2+r} \sup_{x \in I} |v(x) - w(x)|.$$

Since $r > 0$, it follows that T_0 is a contraction mapping on C. The contraction mapping theorem implies that T_0 has a unique fixed-point w in C. To prove existence, extend w to $[0, \infty)$ by setting $w(x) = w_0(x)$ for $x > x_0$. Uniqueness follows from the fact that the restriction to I of any w satisfying the conditions of the Lemma is a fixed-point of T_0.

To prove the rest, let

$$C_1 = \{w \in C : \forall x \in I, w(x) \geq x/r\}$$

$$C_2 = \{w \in C_1 : w \text{ is increasing}\}$$

$$C_3 = \{w \in C_2 : \forall x \leq y \leq x_0, w(y) - w(x) \leq (y - x)/r\}.$$

Note that, for $i \leq 3$, C_i is a complete metric space with respect to the supremum norm. Then, it suffices to show that T_0 maps C_i to itself for $i \leq 3$ (as result follows by the contraction mapping theorem).

To show this for $i = 1$, fix $x \in I$ and $w \in C_1$ and note that

$$T_0 w(x) \geq \frac{x + 2\mathbb{E}_x(s)/r - c_x}{2 + r} \geq \frac{x + 2x/r}{2 + r} = \frac{x}{r}$$

where the first inequality follows from the fact that $w(s) \geq s/r$ for all $s \in I$ and $w_0(s) \geq s/r$ for all $s > x_0$, and the second follows from the fact that $\delta(x) \geq 0$ for $x \leq x_0$.

To show the $i = 2$ case, note that if $w \in C_2$ then $\mathbb{E}_x w \in C_2$. Moreover, $\gamma \in C_2$. Hence $T_0 w \in C_2$.

To show the $i = 3$ case, note that elements of C_3 are Lipschitz continuous, hence absolutely continuous, hence their derivative exists almost everywhere, and it's bounded above by $1/r$. For $y \in I$ such that $y \geq x$,

$$\begin{aligned}
\mathbb{E}_y[\mathbb{I}_{s \leq x_0} w(s)] &= w(x_0) F_y(x_0) - \int_0^{x_0} w'(s) F_y(s) \mathrm{d}s \\
\mathbb{E}_y[\mathbb{I}_{s \leq x_0} w(s)] - \mathbb{E}_x[\mathbb{I}_{s \leq x_0} w(s)] &= w(x_0)[F_y(x_0) - F_x(x_0)] + \int_0^{x_0} w'(s)[F_x(s) - F_y(s)] \mathrm{d}s \\
&\leq w(x_0)[F_y(x_0) - F_x(x_0)] + \frac{1}{r} \int_0^{x_0} F_x(s) - F_y(s) \mathrm{d}s \\
&= [w(x_0) - x_0/r][F_y(x_0) - F_x(x_0)] + \{\mathbb{E}_y[\mathbb{I}_{s \leq x_0} s] - \mathbb{E}_x[\mathbb{I}_{s \leq x_0} s]\}/r. \\
&\leq \{\mathbb{E}_y[\mathbb{I}_{s \leq x_0} s] - \mathbb{E}_x[\mathbb{I}_{s \leq x_0} s]\}/r
\end{aligned}$$

where the first inequality follows from the fact that F_x's are FOSD-ordered w.r.t. x and the fact that the rate of change of w is boudned above by $1/r$, and the second follows from the fact that $w(x_0) \geq x_0/r$. The same reasoning can be used to show that

$$\mathbb{E}_y[\mathbb{I}_{s \geq x_0} w_0(s)] - \mathbb{E}_x[\mathbb{I}_{s \geq x_0} w_0(s)] \leq \{\mathbb{E}_y[\mathbb{I}_{s \geq x_0} s] - \mathbb{E}_x[\mathbb{I}_{s \geq x_0} s]\}/r.$$

Moreover, since δ is decreasing, $y \geq x$ implies that

$$\mathbb{E}_y(s)/r - c_y/2 - (\mathbb{E}_x(s)/r - c_x/2) \leq (y - x)/r$$

Hence

$$
\begin{aligned}
T_0w(y) - T_0w(x) &= \frac{y-x}{2+r} + \frac{2}{2+r}\{\mathbb{E}_y[\mathbb{I}_{s\le x_0}w(s)] - \mathbb{E}_x[\mathbb{I}_{s\le x_0}w(s)] \\
&\quad + \mathbb{E}_y[\mathbb{I}_{s>x_0}w_0(s)] - \mathbb{E}_x[\mathbb{I}_{s>x_0}w_0(s)] - c_y/2 + c_x/2\} \\
&\le \frac{y-x}{2+r} + \frac{2}{2+r}[\mathbb{E}_y(s)/r - c_y/2 - (\mathbb{E}_x(s)/r - c_x/2)] \\
&\le \frac{y-x}{2+r} + \frac{2/r}{2+r}(y-x) \\
&= (y-x)/r.
\end{aligned}
$$

□

Proof of Theorem 1. We prove the result for the baseline example. That is, we assume that $u(a,x) = x(1-a)$. The general proof is available upon request. A Markov strategy profile $\sigma : [0,\infty) \to [0,1]^N$ induces welfare w_σ satisfying

$$
rw_\sigma(x) = x + \sum_{i=1}^{N} \sigma^i(x)\left[\mathbb{E}_x w_\sigma - w_\sigma(x) - \frac{x}{N}\right] \tag{8}
$$

(assuming (8) has a measurable solution). Then (8) is equivalent to

$$
w_\sigma(x) = Tw_\sigma\left[x, \sum_{i=1}^{N}\sigma^i(x)\right]
$$

where T is defined in (6). Let σ be the Markov profile described in Theorem 1. Note that if $x_0 = \hat{x}$ and $w_0(x) = x/r$ then $w_0(x_0) = Tw_0(x_0)$. Then Lemma 1 implies that (8) admits a solution w^* that is continuous, increasing, and such that $w^*(x) \ge x/r$ for all $x \ge 0$.

Hence it suffices to prove that w^* solves the Bellman equation for the welfare benchmark. It is given by

$$
rw(x) = x + \max_{a\in[0,N]} a(\mathbb{E}_x w - w(x) - x/N).^{21} \tag{9}
$$

21. (9) can be derived by noting that $w(x_t) = \max_{a\in[0,N]}(x_t - ax_t/N)\mathrm{d}t + e^{-r\mathrm{d}t}[a\mathrm{d}t\mathbb{E}_{x_t}w + (1 - a\mathrm{d}t)w(x_{t+\mathrm{d}t})]$, approximating $e^{-r\mathrm{d}t} = 1 - r\mathrm{d}t$ and $w(x_{t+\mathrm{d}t}) = w(x_t)$, and neglecting terms in $\mathrm{d}t^2$.

Note that w^* solves (9) on $[\hat{x}, \infty)$. For $x_0 \leq \hat{x}$,

$$\begin{aligned}\mathbb{E}_{x_0} w^* - w^*(x_0) - x_0/N &= \mathbb{E}_{x_0} w^* - \frac{x_0 + N\mathbb{E}_{x_0} w^* - x_0}{N+r} - x_0/2 \\ &= \frac{r\mathbb{E}_{x_0} w^* - rx_0/N - x}{N+r} \geq \frac{\mathbb{E}_{x_0}(s) - rx_0/N - x}{N+r} \geq 0\end{aligned}$$

where the first inequality follows from the fact that $w^*(x) \geq x/r$ for all $x \geq 0$, and the second holds for all $x_0 \leq \hat{x}$. Hence w^* solves (9) for all $x \geq 0$. □

As part of the proof of Theorem 2, I prove the following.

Lemma 2. *In any MPE, at least one player exerts effort if $x_t < \mu/r$, at most one player exerts effort if $x_t = \mu/r$, and no player exerts effort if $x_t > \mu/r$.*

Proof. 'Exert maximal effort at x_0' is strictly better than 'exert no effort' in the single-agent case if $x_0 < \mu/r$. Then this remains true in the two-player game if the opponent plays a Markov strategy and does not exert any effort at x_0. This is because the payoff from 'exert no effort' remains x_0/r whereas the benefit of 'exert maximal effort at x_0' increases weakly. This proves the first part. To prove the rest, fix a MPE and let x' be the supremum of the states at which at least one player exerts some effort. Note that, since the mean of F is finite, any strategy involving some effort in the first period is strictly dominated by 'no effort' if x_0 is sufficiently high. Then $x' < \infty$. *Claim: σ^i: 'exert maximal effort at x_0' is dominated by 'exert no effort' in any game with initial state $x_0 \geq \mu/r$ if opponents play Markov strategies involving no effort above x_0, unless $x_0 = \mu/r$ and all these strategies are 'exert no effort'.* Result follows directly from the claim unless no player exerts effort at state x'. If so, it suffices to prove that $x' \leq \mu/r$. By continuity, it is possible to modify the strategy profile at state x' so that at least one player exerts effort whilst preserving the MPE. Then the claim implies that $x' \leq \mu/r$, and result follows. To prove the claim, fix the Markov strategies and suppose that either $x' > \mu/r$ or the strategies do not all equal 'exert no effort'. Then it suffices to prove that σ^i is not a best-response. Using (4), one can show that, if each opponent were to stop exerting effort when $t = \mathrm{d}t$, then 'exert no effort' would be strictly better than 'exert maximal effort if $t \leq \mathrm{d}t$'.Then, this is also the case if the opponent plays the

Markov strategies. Then, the one-shot deviation principle implies that 'never draw' is a best-response. Hence σ^i is not a best-response. This proves the claim. $\square$

Proof of Theorem 2. I prove existence for the case $u(a,x) = x - c_x a$ for some family $(c_x)_{x \geq 0}$. Setting $c_x = x$ yields the baseline example. The proof of uniqueness, as well as the proof for general payoff functions, is available upon request. The reasoning used to derive (9) implies that, if Player j plays some Markov strategy σ_j, the Bellman equation for Player i is

$$rv_i(x) = x + \max_{a \in [0,1]} a\left(\mathbb{E}_x v_i - v_i(x) - c_x\right) + \sigma^j(x)[\mathbb{E}_x v_i - v_i(x)]. \tag{10}$$

Then, it suffices to find some symmetric, decreasing σ_{M} such that, for some $\hat{y}' < \hat{y}$, $\sigma^i_{\mathrm{M}}(x) = 1$ for $x < \hat{y}'$ and $\sigma_{\mathrm{M}}(x) = 0$ for $x > \hat{y}$, and welfare $w_{\sigma_{\mathrm{M}}}$, defined in (6), solves

$$\mathbb{E}_x w_{\sigma_{\mathrm{M}}} - w_{\sigma_{\mathrm{M}}}(x) - c_x \begin{cases} \geq 0 & \text{if } x < \hat{y}' \\ = 0 & \text{if } x \in [\hat{y}', \hat{y}] \\ \leq 0 & \text{otherwise.} \end{cases}$$

Claim 1. There exists a unique continuous $w_0 : [0,\infty) \to \mathbb{R}$ such that

$$w_0(x) = \begin{cases} x/r & \text{if } x > \hat{y} \\ \mathbb{E}_x w_0 - c_x & \text{otherwise.} \end{cases} \tag{11}$$

Moreover, $w_0(x) \geq x/r$ for all $x \geq 0$ and the rate of change of w_0 is bounded above by $1/r$.

The claim is proved at the end of this proof. Let $\phi(x) = (w_0(x) - x/r)/c_x$, define $\hat{y}' = \max\{0\} \cup \{x \leq \hat{y} : \phi(x) \geq 2\}$, and pick σ_{M} such that $\sigma^i_{\mathrm{M}}(x) = \phi(x)$ for $x \in [\hat{y}', \hat{y}]$. Since $w_0(x) \geq x/r$ for all $x \geq 0$, then σ is well-defined. Moreover, σ is decreasing on $[\hat{y}', \hat{y}]$ since the rate of change of w_0 is bounded above by $1/r$. Then, it is clear from (6) that $w_0(x) = w_{\sigma_{\mathrm{M}}}(x)$ for all $x \geq \hat{y}'$. If $\hat{y}' = 0$, we are done. Otherwise, since w_0 is continuous, $w_0(x_0) = Tw_0(x_0, 2)$ if $x_0 = \hat{y}'$. Then, from Lemma 1, there exists a unique $w : [0,\infty) \to \mathbb{R}$ such that $w(x) = w_0(x)$ for $x \geq \hat{y}'$ and $w(x) = Tw(x,2)$ for $x \leq \hat{y}'$. Moreover, w is continuous and $w(x) \geq x/r$

for all $x \geq 0$. Since $\sigma^i_{\mathrm{M}}(x) = 1$ for all $x \leq \hat{y}'$ and w is continous, it follows that $w = w_{\sigma_{\mathrm{M}}}$. Finally, note that, for $x \leq \hat{y}'$, Note that

$$\begin{aligned}
\mathbb{E}_x w - w(x) - c_x &= \mathbb{E}_x w - \frac{x + 2\mathbb{E}_x w - c_x}{2+r} - c_x \\
&= \frac{(r\mathbb{E}_x w - x/r - rc_x)}{2+r} \\
&\geq \frac{\mathbb{E}_x(s) - x - rc_x)}{2+r} \\
&\geq 0
\end{aligned}$$

where the first inequality follows from the fact that $w(x) \geq x/r$ for all $x \geq 0$, and the second follows from the fact that $\mathbb{E}_x(s) - x - rc_x$ is decreasing (since δ is decreasing and c_x is increasing) and equals 0 at $x = \hat{y}$. □

Proof of Claim 1. Let A be the set of all $x \geq 0$ such that there exists a unique continuous $w_0 : [x, \infty) \to \mathbb{R}$ satisfying (11). Clearly $\hat{y} \in A$, so A is not empty. Let $x_0 = \inf A$. Then it suffices to prove that $x_0 = 0$. Suppose $x_0 > 0$ and seek a contradiction. Use the contraction mapping theorem. □

Proof of Theorem 3. Given a PPE σ and $x \geq 0$, let $w_\sigma(x)$ be the supremum of welfares induced by σ, across all public histories leading to stock x.[22] Let C be the set of functions $w : [0, \infty) \to \mathbb{R}$ such that $w(x) \in [x/r, w^*(x)]$ for all $x \geq 0$, and $w(x) = x/r$ if $x > \hat{y}$. I claim that $w_\sigma \in C$ for any PPE σ. Clearly, for any $x \geq 0$, a player exerting no effort obtains a payoff of at least x/r in any game with initial state x, hence $w_\sigma(x) \geq x/r$ for all $x \geq 0$. Moreover, by definition of w^*, $w_\sigma(x) \leq w^*(x)$ for all $x \geq 0$. Finally, from Lemma 2, $w_\sigma(x) = x/r$ for all $x > \hat{y}$.

For $w \in C$, let $\bar{T}w : [0, \infty) \to \mathbb{R}$ be such that $\bar{T}w(x) = x/r$ for $x > \hat{y}$, and

$$\bar{T}w(x) = \max\{Tw(x, 1), \min\{\mathbb{E}_x w - c_x, Tw(x, 2)\}\}$$

for all $x \in I$. Since w^* is continuous on $[0, \hat{y}]$, then C is a complete metric space with respect to the supremum norm. Moreover, for any $x' \in I$, $Tw(x', \cdot)$ is increasing since $w(x) \geq x/r$ for all $x \geq 0$. Hence $x'/r \leq Tw(x', 1) \leq \bar{T}w(x') \leq Tw(x', 2) \leq w^*(x')$. Therefore it is clear that $\bar{T}$ maps C to C, Finally, $\bar{T}$ satisfies Blackwell's

22. Note that w_σ solves (6) when σ is a MPE.

sufficient conditions for a contraction mapping. Namely, if $w, v \in C$ are such that $w(x) \geq v(x)$ for all $x \geq 0$, then $Tw(x,a) \geq Tv(x,a)$ for $a \in \{1,2\}$, and $\mathbb{E}_x w \geq \mathbb{E}_x v$, hence $\bar{T}w(x) \geq \bar{T}v(x)$. Moreover if, for some $a \geq 0$, $v(x) = w(x) + a$ for all $x \in I$, then $\bar{T}v(x) \leq Tv(x,2) = \bar{T}w(x) + \beta a$ for all $x \geq 0$, where $\beta = 2/(r+2) \in (0,1)$. Hence $\bar{T}$ is a contraction mapping on C, so it has a unique fixed-point in C.

Let C_0 be the set of $w \in C$ such that $w(x) \leq \bar{T}w(x)$ for all $x \in I$. I claim that $w_\sigma(x) \in C_0$. Fix $x \in I$. Since $w_\sigma(x') \geq x'/r$ for all $x' \geq 0$, then $w_\sigma(x) \leq Tw(x,2)$. By definition of $w_\sigma(x)$, $w_\sigma(x) = \lim_n w_\sigma(h^n)$ for some sequence $(h_n)_{n=1}^\infty$ of public histories leading to stock x. Suppose that, for some $T > 0$, there exists a subsequence $(h'_n)_{n=1}^\infty$ such that, for any $n = 1, 2, \ldots$, conditional on the stock not increasing before time T, at most one player exerts effort at any time $t \leq T$. If $\sigma^i(x) = 0$ for some $i \leq 2$, then $w_\sigma(x) \leq Tw(x,1)$. Otherwise, both players are willing to exert effort at x. Then (10) implies that, for $i = 1,2$, $\mathbb{E}_x v_i - v_i(x) \geq c_x$, where v_i is the value function of Player i. Averaging across i implies that $w_\sigma \leq \mathbb{E}_x w_\sigma - c_x$. This proves the claim.

Let $\bar{w} : [0,\infty) \to \mathbb{R}$ be such that $\bar{w}(x) = \sup_{w \in C_0} w(x)$ for all $x \geq 0$. Clearly, $\bar{w} \in C_0$. Moreover, it is clear that $\bar{w}(x) \geq \bar{T}\bar{w}(x)$ for all $x \in I$ since, for any $x \in I$, there exists $w \in C_0$ such that $w = \bar{w}$ on (x,∞) and $w(x) = \bar{T}\bar{w}(x)$. Hence $\bar{w}$ is a fixed-point of $\bar{T}$.

I claim that $\bar{w} = w_{\sigma_{\mathrm{P}}}$ for some Markov profile σ_{P} with the stated properties. Let $w_0 \in C$ be such that $w_0(x) = Tw_0(x,1)$ for all $x \in I$. If w_0 is a fixed-point of $\bar{T}$, then the claim holds with $\hat{z} = 0$. Otherwise, let $\hat{z} = \sup\{x \in I : \bar{T}w_0(x) > w_0(x)\}$, and let $w_1 \in C$ be such that

$$w_1(x) = \begin{cases} \mathbb{E}_x w_1 - c_x & \text{if } x \leq \hat{z} \\ w_0(x) & \text{otherwise.} \end{cases}$$

The rest of the argument to prove the claim follows the steps of Proposition 2.

It remains to show that σ_{P} is an MPE. The argument used to construct the most efficient pure-strategy MPE in Theorem **??** can be adapted to construct a Markov profile σ_{N} such that (a) $\sigma_{\mathrm{N}}^i(x) = 0$ for $i = 1,2$ if $x > \hat{y}$, (b) $\{\sigma^i(x) : i = 1,2\} = \{0,1\}$ if $x \in (\hat{z}, \hat{y}]$ and (c) $\sigma^i(\hat{z}) = 1/2$ for $i = 1,2$ and expected discounted payoffs are symmetric at $x_0 = \hat{z}$.

Now, let $\sigma_{\mathrm{P}} : [0, \infty) \to [0, 1]$ be such that, for $i = 1, 2$

$$\sigma^i_{\mathrm{P}} = \begin{cases} \sigma^i_{\mathrm{N}}(x) & \text{if } x > \hat{z} \\ (\bar{w}(x) - x/r)/c_x & \text{if } \hat{z}' < x \leq \hat{z} \\ 1 & \text{otherwise} \end{cases}$$

One can verify that $\bar{w}$ is the welfare induced by σ_{P}. It remains to prove that σ_{P} is an equilibrium for all $x_0 \leq \hat{z}$. It is clear that σ_{P} is an equilibrium for $x_0 > \hat{z}$. Moreover, σ_{P} is symmetric for $x_0 \leq \hat{z}$. Since payoffs are symmetric at $x_0 = \hat{z}$, it follows that they are symmetric at all $x_0 \leq \hat{z}$. Then, the definition of $\bar{w}$ implies that σ_{P} is an equilibrium. □

Proof of Theorem 4. The reasoning is very similar to the proof of Theorem 3. Welfare is maximal if the agent inducing in increment recieves the highest possible equilibrium continuation payoff. □