Working papers series

WP ECON 21.15

Optimal Management of Evolving Hierarchies

Jens L. Hougaard
Department of Food and Resource Economics, University of Copenhagen

Juan D. Moreno-Ternero
Department of Economics, Universidad Pablo de Olavide

Lars P. Osterdal
Department of Economics, Copenhagen Business School

Keywords: Optimal allocation schemes; Hierarchies; Management; Nash equilibrium; Blockchain.

JEL Classification: C70, L24, M52 .

Department of Economics

Optimal Management of Evolving Hierarchies*

Jens Leth Hougaard[†‡]

Juan D. Moreno-Ternero[§]

Lars Peter Østerdal[¶]

August 3, 2021

Abstract

We study the optimal management of evolving hierarchies of revenue-generating agents. The initiator invests into expanding the hierarchy by adding another agent, who will bring revenues to the joint venture and who will invest herself into expanding the hierarchy further, and so on. The higher the investments (which are private information), the higher the probability of expanding the hierarchy. An allocation scheme specifies how revenues are distributed, as the hierarchy evolves. We obtain schemes that are socially optimal and initiator-optimal respectively. Our results have potential applications for blockchain, cryptocurrencies, social mobilization and multi-level marketing.

JEL Classification: C70, L24, M52

Keywords: Optimal allocation schemes; Hierarchies; Management; Nash equilibrium; Blockchain

*We are grateful to Manel Baucells (department editor of this journal), the anonymous associate editor, and three anonymous referees for extremely helpful comments and suggestions. We also thank Jens Gudmundsson, Xiaoye Liao, Herve Moulin, Maxwell Pak, Takuro Yamashita and audiences at National University of Singapore, University of Southern Denmark (GEM workshop), Shanghai University of Finance and Economics, Korea University (Int. Conf. on Microeconomics), the 2019 SAET Meeting (Ischia), the 2019 Stony Brook Game Theory festival, NYU Shanghai, NYU Abu Dhabi, Copenhagen Business School, and University of Copenhagen (workshop on Blockchains and Economic Design), for valuable comments and suggestions on earlier versions of this paper. Financial support from the Independent Research Fund Denmark (DFF-6109-000132) and the Spanish Ministry of Economy and Competitiveness (ECO2017-83069-P, PID2020-115011GB-I00) is gratefully acknowledged. Hougaard is furthermore grateful for financial support from the Center for Blockchains and Electronic Markets funded by the Carlsberg Foundation (CF18-1112).

†Department of Food and Resource Economics, University of Copenhagen, Denmark.

‡Economics, NYU Shanghai, China.

§Department of Economics, Universidad Pablo de Olavide, Spain.

¶Department of Economics, Copenhagen Business School, Denmark.

1 Introduction

It has long been argued that the basic structure (or order) of the world is a hierarchical structure (e.g., Salthe, 1985). Hierarchies seem to be a pervasive feature of the organization of natural and artificial systems (e.g., Corominas-Murtra et al., 2013), and hierarchical systems are frequent, for instance, in large scale industrial automation (Mesarovic et al., 1970), clinical trials (e.g., Azoulay, 2004), decentralized computer networks (e.g., Furth, 2010) or decision making and planning (e.g., Saaty, 1986; Harker and Vargas, 1987). Hierarchies influenced the survival of prisoners of war during World War II (e.g., Holderness and Pontiff, 2012). Social hierarchies influence primate health (e.g., Sapolsky, 2005). Top-down (hierarchical) mechanisms are implemented for memory retrieval and assimilation (e.g., Rajasethupathy et al., 2015). And even the brain could be modeled as a hierarchical system (e.g., Brocas and Carrillo, 2008).

In this paper, we study how to manage incentives in an evolving hierarchy of agents. In our stylized model, the first agent (the initiator) invests resources in order to add a second (revenue-generating) agent to the hierarchy. This second agent, if added, invests herself in adding a third (revenue-generating) agent, and so on. The more an agent invests, the higher the probability that she adds another agent. An agent's incentive to invest is determined by her expected return, comprising what is potentially obtained from the next agent, and from further agents added later to the hierarchy. Thus, a transfer of revenues generated by lower-ranked (i.e., newer) agents to higher-ranked (i.e., older) agents is essential for incentivizing investments (and, thus, potentially expanding the hierarchy). Our main research question is how this process can be optimally managed through appropriately allocating the revenues obtained in the hierarchy.

Potential areas of application include mining in blockchains (Schrijvers et al., 2017), social mobilization (Pickard et al., 2011), and multi-level marketing (Emek et al., 2011). In the first case, miners perform (costly) computations as part of the "proof-of-work" protocol in order to add another block to the blockchain. If a miner is successful solving the cryptographic puzzle of the hash functions, and a new block is added, a reward is paid. The issue is to optimally allocate this reward among the successful miner and her predecessors. In the case of social mobilization, a task has to be solved and an initiator exerts effort in recruiting (mobilizing) agents to help solving this task, who exert effort themselves to recruit more agents, and so on. Mobilizing agents increases the probability of solving the task, thus adding value to the recruitment hierarchy. In multi-level marketing, the initiator can generate a revenue and recruit a sales partner, that also generate a revenue and recruit another sales partner, and so on.

1.1 Our Results

In our model, a planner, or system designer, could aim at maximizing the expected payoff of the initiator (in particular, if the initiator is the planner herself), but could also consider a social objective and maximize the overall expected value of the hierarchy. Our first result (Theorem 1) actually shows that a first-best socially optimal investment profile exists, and is unique. Moreover, as agents are identical (except, obviously, by their position in the hierarchy) so is the optimal investment of each agent.

We then move to a decentralized framework where each allocation scheme induces a game in which agents choose their investment level strategically. The induced game is *supermodular* (i.e., the marginal value of one player's action is increasing in the other players' actions) so there always exists at least one Nash equilibrium in pure strategies. We consider two natural optimality notions for allocation schemes: an allocation scheme is *decentralized socially optimal* if it induces investments yielding the highest possible overall expected value of the hierarchy; an allocation scheme is *initiator-optimal* if it induces investments yielding the highest possible expected payoff for the initiator.

Concerning decentralized social optimality, our Theorem 2 shows that there is a unique socially optimal allocation scheme characterized by imposing full transfers to the immediate predecessors. In other words, transferring the full value of recruiting a new agent for the hierarchy to the immediate predecessor gives maximal incentive to invest and thereby maximizes the expected length of the hierarchy. The corresponding equilibrium investment profile is constant (i.e., agents invest the same amount), and leads to under-investment compared to the first-best social optimum, as shown in Proposition 1.

Concerning initiator-optimality, we show that there exist multiple optimal schemes, but we focus on a particularly interesting one, which comes in the form of a one-parameter family. It imposes to each member a full transfer to only two recipients: the immediate predecessor and the initiator. Furthermore, a constant ratio among the transfers is imposed for all members. More precisely, there exists a parameter $\alpha \in (0,1)$ such that each agent transfers $100\alpha\%$ of her revenue to her immediate predecessor and the residual percentage $100(1-\alpha)\%$ to the initiator. This scheme is remarkably simple. Another advantage of it, as shown in the proof Theorem 3, is that all agents (except for the initiator) have a dominant strategy in the induced game. As mentioned above, it is not uniquely determined though. We can actually find other transfers yielding the same equilibrium investment profile and expected payoff to the initiator. In fact, the main characteristic of an initiator-optimal allocation scheme is that it yields the same expected payoff to all agents except for the initiator. We exemplify this feature with a reinterpretation of the class of bubbling-up (or geometric) schemes characterized in Hougaard et al., (2017), as described in Proposition 2. Our result confirms and elaborates on the claim made in Pickard et al (2011), that using a bubbling-up scheme to reallocate a reward provides optimal incentives for mobilization. In our setting, we actually show that a planner can actually choose optimal allocation schemes from a much broader class, including the simpler one-parameter family from Theorem 3.

A remarkable feature, shared by all of our results described above, is that we obtain clean comparative statics for a general specification of probability functions, which is surprising. In other words, we do not need to specify a particular probability function to obtain our results.

1.2 Related literature

Our model of managing evolving hierarchies is somewhat related to a specific network intervention known in the literature as *induction*. The term *network interventions* describes the process of using social network data to accelerate behavior change or improve organizational performance (e.g., Valente, 2012). Induction interventions stimulate peer-to-peer interaction to create cascades in behavioral diffusion, implicitly endorsing that secondary incentives can be more efficient and effective than primary incentives, at least in some contexts. For instance, media marketing campaigns often rely on word-of mouth strategies, such as encouraging users to recommend products to their connections, who would do the same themselves (e.g., Aral and Walker, 2011). In *respondent-driven sampling* (e.g., Heckathorn, 1997), a form of chain-referral sampling also known as "snowball methods", individuals recruit others to receive an intervention, who subsequently encourage additional people to participate, and so on.[1]

Our analysis focusses on the dichotomy between centralized and decentralized management of hierarchies. In that sense, we are close to a literature dealing with search and organizational hierarchy, in which it has been argued that a hierarchy with a central decision maker at the top can speed up problem solving, but possibly at the cost of solution quality compared with results of a decentralized search (e.g., Rivkin, 2000; Mihm et al., 2010).

Regarding the decentralized management of hierarchies, our analysis of optimal allocation

[1] Goldlücke (2017) also deals with the strategic recruiting in ongoing hierarchies. Her emphasis is on analyzing the effect of skill-based promotions on incentives to recruit, exploring how the management rule "A's hire A's and B's hire C's" can make sense in a game-theoretic model.

schemes is somewhat reminiscent of the analysis in Galeotti et al., (2019), who design and evaluate welfare-enhancing tax policy schemes in complex supply chains consisting of primary and final good producers.[2]

From a different vantage point, there has been a growing interest in the literature dealing with resource allocation in the presence of a hierarchical structure (see, for instance, Hougaard (2018) and the references cited therein). This interest can be traced back to Claus and Kleitman (1973) and Bird (1976), with their canonical cost sharing problem within a rooted tree, or Littlechild and Owen (1973) with the so-called *airport problem*, in which the runway cost has to be shared among different types of airplanes with a linear graph representing the runway. More recently, Hougaard et al., (2017) consider the problem of distributing the proceeds generated from a joint venture in which the participating agents are hierarchically organized.[3] They characterize a family of allocation rules where revenue 'bubbles up' in the hierarchy, ranging from the *no-transfer* rule (where no revenue bubbles up) to the *full-transfer* rule (where all the revenues bubble up to the top of the hierarchy). Intermediate rules within that family are reminiscent of popular incentive mechanisms for social mobilization or multi-level marketing (e.g., Pickard et al., 2011; Emek et al., 2011) and can also be seen as specific *geometric (incentive tree) mechanisms* (e.g., Lv and Moscibroda, 2013) that are usually considered in the computer science literature.

In Juarez et al., (2020), agents are endowed with time, which is invested in projects that generate profit. A mechanism divides the profit depending on the allocation of time, as well as the amount of profit made by every project. They characterize the mechanisms that incentivize agents to contribute their time to a level that results in the maximal aggregate profit at the Nash equilibrium, regardless of the production functions involved. Within this class, they also characterize the mechanisms that are monotone on the addition of time to agents as well as on the payoffs of the agents with respect to technological improvements in the generation of profit. In contrast with our work, they focus on a static case, whereas we deal with a dynamic extension when the network is not given but generated based on the sharing rule.

[2]In a similar vein, Bimpikis et al., (2019) consider a model of competition among firms that produce a homogeneous good in a networked environment, in which a bipartite graph determines which subset of markets a firm can supply to.

[3]This is reminiscent of the problem of sharing a polluted river (e.g., Ni and Wang, 2007; Dong et al., 2012), with the modification of considering negative revenues, and thus interpreting them as costs. Another recent and somewhat related resource allocation problem under the presence of hierarchical structures is Juarez et al., (2018).

Our model could also be considered as a moral hazard optimal contracting problem, as pioneered by Holmstrom (1982), who showed that there is no allocation rule that induces a game with first-best investment profile in equilibrium. Our contracting problem is a dynamic specification of Holmstrom's problem with the innovative aspect that the joint monetary outcome is an asymmetric function of the agents' (subordinates') efforts, due to the hierarchical structure. The fact that transfers must meet the budget constraint (revenue in this case) implies that each agent cannot be rewarded for the full marginal profit that her effort generates. Moreover, the informativeness principle (a signal is informative about effort if and only if it has value for contracting) would indicate that second-best incentive contracts should condition pay on the measures that are best indications of each agent's effort. This is, essentially, what drives our Theorems 1, 2 and Proposition 1. Our analysis of initiator-optimal contracts could be interpreted as a study of how much of the chain's value can one party appropriate when limited by revenue-sharing contracts.

Within the scope of the literature on teams, also initiated with Holmstrom (1982), Strausz (1999) considers sequential partnerships and shows that there exists a balanced budget-sharing rule that achieves social efficiency. Winter (2006) also studies optimal incentive schemes in organizations where agents perform their tasks sequentially. In his model, agents' effort decisions are mapped into the probability of the project's success. He characterizes the unique optimal investment-inducing mechanism, i.e., rewards are allocated among agents so as to induce all of them to exert effort in equilibrium at minimal cost to the principal. A crucial feature of such a mechanism is that agents whose efforts are relatively unobservable by their peers need to be promised higher rewards for a successful outcome to make them exert effort.[4]

Finally, our work is also related to a new literature dealing with blockchain, the fundamental technology underlying the emergence of cryptocurrencies such as Bitcoin (e.g., Cheng et al., 2019, Huberman et al., 2019; Azevedo et al., 2020). As outlined above, in proof-of-work protocols, miners take on the computational work required to assemble new blocks and commit them to the shared ledger. Each time a miner commits a new block to the chain it can assign a predefined amount of the crypto token to itself as a reward. This reward is combined with the transactions fees participants may have included in their individual transactions to incentivize miners (to prioritize them over others in the construction of the next block) (e.g., Catalini and Gans, 2020). Our results can shed light on the design of such optimal incentive schemes.

[4]More strikingly, such a discrimination among equals is also a feature of optimal investment-inducing mechanism in the counterpart model in which agents exert effort decisions simultaneously (e.g., Winter, 2004).

2 The model and optimality notions

We imagine a dynamic process where agents (or tasks) can be added (performed) in sequence. The probability of successfully adding an agent (solving a task) depends on the amount of resources invested in the process. In case of a successful addition, a revenue is obtained. By reallocating these revenues, through a so-called allocation scheme, agents can be incentivized to invest in the process.

Formally, the set of potential agents is identified with the natural numbers including 0, i.e., $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. Agent 0 represents the initial agent: to be thought of as the initiator, designer, boss, patent owner, patriarch, etc., depending on the application in mind. Agent 0 starts the dynamic process by trying to attract agent 1 as a follower. The probability of succeeding is increasing in the amount of resources that agent 0 invests in the process. If agent 0 succeeds, this follower (agent 1) can now start the same process (as her predecessor) to attract agent 2 as a follower, and so forth, until some agent (n) is unsuccessful in getting a follower. The outcome of the process is then a realized hierarchy, representing an ordering of agents with higher numbers indicating a lower rank.

Let $x_i \in \mathbb{R}_+$, denote the amount of resources that agent i invests in the process. We assume that the ability of adding a follower is the same for every agent and is formalized by a function $p : \mathbb{R}_+ \to [0, 1)$, which assigns for each investment level $x_i \in \mathbb{R}_+$, the probability $p(x_i) \in [0, 1)$ that the investment is successful. We refer to p as the *technology* and assume that it is a strictly increasing, differentiable and strictly concave function, satisfying that $p(0) = 0$, $\lim_{x_i \to 0^+} p'(x_i) = +\infty$, and such that $\frac{p(x_i)}{p'(x_i)}$ is convex at $(0, +\infty)$.[5] Note that no matter how much a given agent invests there is no guarantee of getting a follower. That is, $p(x_i) < 1$ for all $x_i \in \mathbb{R}_+$, which will be crucial for the ensuing analysis.

Also, we assume that agents' investment decisions are private information and that there is no budget constraint.

When an agent joins the hierarchy a value is added. To simplify the presentation, we assume that each agent joining the hierarchy adds the same value (e.g., provides the same revenue),

[5]Although crucial, we believe these assumptions on technologies are natural. In particular, the latter assumption states that the inverse semi-elasticity is convex. Thus, it is a regularity condition ensuring that $p'(x_i)$ diminishes relative to $p(x_i)$ at an increasing rate (note that $\frac{p(x_i)}{p'(x_i)}$ is increasing), which therefore rules out technologies with "soft" kinks. A broad range of relevant technologies satisfy all conditions; for instance, the entire family given by $p(x_i) = \frac{x_i^\beta}{\theta + x_i^\beta}$ where $\beta \in (0, 1)$ and $\theta > 0$.

which we normalize to 1.[6]

In line with the rank of the hierarchy, we assume that the value of a realized agent can only be transferred upwards in the hierarchy (that is, from agents with higher numbers to agents with lower numbers) or remain with that agent. The size of these transfers is fixed by a predefined allocation scheme known by all agents ex-ante. Thus, the (normalized) value of 1 is fully allocated among the realized agent and her predecessors.[7]

Formally, an *allocation scheme* $t = \{t_{ij}\}_{\{ij\}\subset \mathbb{N}_0}$ is given by transfers $t_{ij} \in [0,1]$ such that $t_{ij} = 0$ for each $i < j$ and $\sum_{j\leq i} t_{ij} = 1$ for each i. Note that t_{ij} is interpreted as the amount that agent i allocates to agent j, provided i joins the hierarchy.

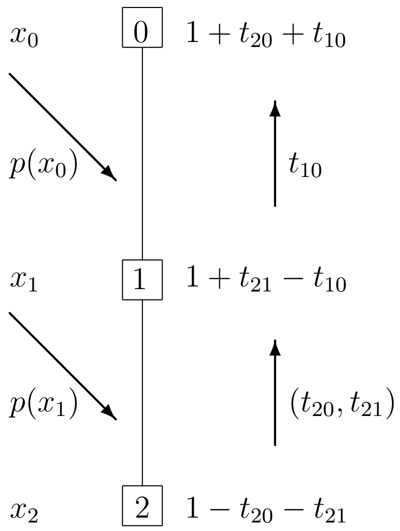

Figure 1: A realized hierarchy with three agents (the initiator and two successors).

For each agent $i \in \mathbb{N}_0$, each allocation scheme t, and each profile of investment choices $x = (x_0, x_1, x_2, \dots)$, let $\mathbb{E}_i(t, x)$ denote the expected payoff for agent i conditional on the realization of that agent.[8] Formally,

$$\mathbb{E}_i(t,x) = t_{ii} - x_i + p(x_i)t_{(i+1)i} + p(x_i)p(x_{i+1})t_{(i+2)i} + \cdots = t_{ii} - x_i + \sum_{l=1}^{+\infty}\prod_{k=0}^{l-1} p(x_{i+k})t_{(i+l)i}. \quad (1)$$

For each given profile of investments, the *social value* is given by the expected total payoff. Given a profile of investments x, the allocation scheme t is irrelevant for the computation of

[6]Relaxing this assumption would open the door to explore different features, such as the selective recruiting of individuals (certainly an interesting aspect, but beyond the scope of this paper).

[7]Instead of assuming this, we could consider that the (normalized) value of 1 is fully allocated among predecessors, excluding the realized agent. The results we shall present would also hold with this alternative definition.

[8]Note that infinite expected payoffs are not ruled out from the outset.

the social value because transfers among agents cancel out. Therefore, it can be written as

$$\begin{aligned}\mathbb{V}(x) &= \mathbb{E}_0(t,x) + p(x_0)\mathbb{E}_1(t,x) + p(x_0)p(x_1)\mathbb{E}_2(t,x) + \dots \\ &= (1-x_0) + p(x_0)\,(1-x_1) + p(x_0)p(x_1)\,(1-x_2) + \dots \\ &= (1-x_0) + \sum_{i=1}^{+\infty}\prod_{k=0}^{i-1} p(x_k)\,(1-x_i)\,.\end{aligned}$$

For each given profile of investments, x, we define the *expected length* of the hierarchy (which could be interpreted as the expected total gross return from the hierarchy) as

$$\mathbb{L}(x) = 1 + p(x_0) + p(x_0)p(x_1) + \dots = 1 + \sum_{i=1}^{+\infty}\prod_{k=0}^{i-1} p(x_k),$$

and the *expected investment* as

$$\mathbb{I}(x) = x_0 + p(x_0)x_1 + p(x_0)p(x_1)x_2 + \dots = x_0 + \sum_{i=1}^{+\infty}\prod_{k=0}^{i-1} p(x_k)x_i.$$

Note that, provided $\mathbb{V}(x)$, $\mathbb{L}(x)$ and $\mathbb{I}(x)$ are all finite, then $\mathbb{V}(x) = \mathbb{L}(x) - \mathbb{I}(x)$.

We say that a profile of investments x^* is *first-best socially optimal*, if the social value is maximized. As transfers are irrelevant, this kind of profile can be equivalently obtained analyzing a single-agent dynamic investment problem.

Note that *first-best socially optimal* profiles are *dynamically consistent*. By this we mean that they also maximize *downward social value* along the hierarchy. Formally, at each period l, the downward social value is given by

$$\mathbb{V}^l(x_l, x_{l+1}, \dots) = (1-x_l) + \sum_{i=l+1}^{+\infty}\prod_{k=l}^{i-1} p(x_k)\,(1-x_i)\,. \qquad (2)$$

Then, if $x^* = (x_0^*, x_1^*, x_2^*, \dots)$ is *first-best socially optimal* it follows that $x^{l*} = (x_l^*, x_{l+1}^*, \dots)$ maximizes (2).

In a decentralized context, each allocation scheme induces a game in which agents decide their investment choices (which are private information). A profile of investments $x^* = (x_0^*, x_1^*, \dots)$ is a (Nash) *equilibrium* of this game if, for each $i \in \mathbb{N}_0$,

$$\mathbb{E}_i(t, (x_i^*, x_{-i}^*)) \geq \mathbb{E}_i(t, (x_i, x_{-i}^*)),$$

for each $x_i \in \mathbb{R}_+$.[9]

[9] Note that this is not an "open loop" equilibrium notion, as a predecessor's level of investment cannot affect an agent's optimal choice.

In equilibrium, each agent $i \in \mathbb{N}_0$, solves the following first order condition, if $x_i > 0$:

$$p'(x_i) = \frac{1}{t_{(i+1)i} + \sum_{l=2}^{+\infty} \prod_{k=1}^{l-1} p(x_{i+k}) t_{(i+l)i}}. \tag{3}$$

Note that (3) shows that every agent has a unique best response given the other agents' investment levels (equilibrium or not). In fact, the game is supermodular (e.g., Topkis, 1979) as there are monotonically increasing best replies.[10]

As we shall argue later, agents' investments are, de facto, bounded and, thus, the fact that the game is supermodular means that for any technology p, and any allocation scheme t, there exists a pure strategy Nash equilibrium.

Remark 1 *Uniqueness of equilibrium is not guaranteed. To show this, consider the technology defined by*

$$p(x_i) = \frac{\sqrt{x_i}}{\frac{1}{4} + \sqrt{x_i}}.$$

Let t be the allocation scheme such that $t_{00} = t_{11} = 1 = t_{i(i-2)}$, for each $i \geq 2$, and $t_{ij} = 0$ otherwise. As $p(0) = 0$, it follows that the profile $(0, 0, \dots)$ is an equilibrium. Now, (3) corresponds to

$$1/8 = (1/4 + \sqrt{x_i})^3,$$

which has the solution $x_i^* = 1/16$, with $p(x_i^*) = 1/2$. Thus, the profile $(x_i^*, x_i^*, \dots)$ is also an equilibrium.

In a decentralized context, there are two natural optimality notions: one from the viewpoint of a social planner; and one from the viewpoint of the initiator.

We say that an allocation scheme t^* is *decentralized socially optimal* if there exists a profile of investments x^*, which constitutes an equilibrium for the game induced by t^*, and $\mathbb{V}(\bar{t}, \bar{x}) \leq \mathbb{V}(t^*, x^*)$, for each allocation scheme $\bar{t}$, inducing a game for which $\bar{x}$ is an equilibrium. Such a supporting equilibrium will sometimes be called *socially optimal equilibrium*.

We say that an allocation scheme, t^*, is *initiator-optimal* if there exists a profile of investments x^*, which constitutes an equilibrium for the game induced by t^*, and $\mathbb{E}_0(\bar{t}, \bar{x}) \leq \mathbb{E}_0(t^*, x^*)$, for each allocation scheme $\bar{t}$, inducing a game for which $\bar{x}$ is an equilibrium. Such a supporting equilibrium will sometimes be called *initiator-optimal equilibrium*.

[10] Note that increasing x_j, so does $p(x_j)$. Thus, the denominator at (3) increases and the fraction decreases. By strict concavity of $p(\cdot)$, the optimal x_i increases.

In other words, in a decentralized context, the socially optimal scheme maximizes the expected payoff of the hierarchy as a whole, while an initiator-optimal allocation scheme maximizes the expected payoff of the initiator. The latter can be seen as a particularly natural option as the equilibrium of the game where the initiator sets the sharing rule, aiming to maximize her payoff.

3 The results

We provide in this section the results of our paper. For a smooth passage, we defer their proofs to the Appendix. We start out analyzing the centralized approach to our model. Then, we turn to the decentralized approach, concentrating first on socially optimal allocation schemes and on initiator-optimal allocation schemes afterwards.

3.1 Centralized social optimality

Our first result states the existence and uniqueness of a first-best socially optimal investment profile, as well as some of its properties.

We say that a profile of investments x is *constant* if $x_i = x_j$ for each pair $i, j \in \mathbb{N}_0$. Then, we have the following result.

Theorem 1 *There exists a unique first-best socially optimal investment profile. It is constant.*

Theorem 1 states that the socially optimal investment profile is unique and constant among members of the hierarchy. As mentioned above, the complete proof of Theorem 1 appears in the Appendix. Let us simply note here that *existence* relies on the fact that the strategy space of each agent can be restricted to the interval $[0, 1]$ (a compact set). By Tychonoff's theorem, which states that the product of any collection of compact topological spaces is compact with respect to the product topology, it follows that the (restricted) domain of the social value function ($\mathbb{V}$) is a compact set. As $\mathbb{V}$ is a continuous real-valued function therein, the extreme value theorem concludes. As for *uniqueness*, it essentially follows from the property of *dynamic consistency* and the fact that two constant profiles cannot achieve the (same) maximum social value.

Theorem 1 sets the benchmark case, before moving to the decentralized context analyzed next, which is the main part of this work. Our main concern is to analyze the design of optimal transfer schemes for the management of evolving hierarchies. Theorem 1 will nevertheless be useful, for instance, to compute the price of anarchy, when moving from the centralized to the decentralized context.

3.2 Decentralized social optimality

Our second result states that the (unique) socially optimal allocation scheme imposes full transfers to the immediate predecessors.

Theorem 2 *The allocation scheme* $t^* = \{t^*_{ij}\}_{\{ij\}\subset\mathbb{N}_0}$, *where*

$$t^*_{ij} = \begin{cases} 1 & \text{if } i = j+1 \geq 1, \text{ or } i = j = 0 \\ 0 & \text{otherwise.} \end{cases}$$

is the unique decentralized socially optimal scheme.

The proof of Theorem 2, which appears in the Appendix, relies on a crucial step showing that, for each non-constant equilibrium profile, there exists a constant equilibrium profile with the same expected length from the hierarchy, and higher social value. This mostly follows from the fact that the maximal expected length is a strictly concave function of the expected investment. As a decentralized socially optimal allocation scheme involves null self-transfers, we obtain from the above that the scheme t^* is optimal. Uniqueness follows the fact that the equilibrium profile is constant.

It is interesting (and somewhat surprising) to obtain that such a simple scheme turns out to be the unique decentralized socially optimal scheme. If one interprets that agents exert effort recruiting, then imposing full transfers to immediate predecessors can actually be seen as a way of rewarding effort, thus being a *fair* scheme. Furthermore, the scheme exhibits an interesting feature that can be considered as some sort of *impartiality*, a basic requirement of justice (e.g., Moreno-Ternero and Roemer, 2006). Namely, conditional on the fact that an agent exists (i.e., she is recruited), then all agents get the same in expectations. Now, if one instead interprets that initial agents (especially, the initiator herself) bring some additional value (beyond the above-mentioned effort exerted while recruiting) to the joint venture, then one might expect that other schemes (rewarding them further) would arise. But even under that interpretation (which is not properly formalized in our model), we would not consider the scheme from Theorem 2 as *unfair*.

Theorem 2 can actually be viewed as formal support for the optimality of mining reward (*coinbase*) transactions in Bitcoin where the full reward goes solely to the successful miner. Indeed, it seems reasonable to argue that the original objective in Nakamoto (2008) was to define a (decentralized) socially optimal mining reward scheme in order to ensure long run system sustainability. The main message from Theorem 2 is that social optimality is obtained when the entire reward from creating a new block goes to the agent that invests effort (computational power) in authorizing the block. In our model, all miners are considered as one agent and each individual miners' expected reward is proportional to her computational power. Viewed in isolation, this proportionality rests on firm normative ground (e.g., Chen et al., 2019).

Note that the socially optimal scheme implies that some realized agent will end up getting a negative payoff (by her unsuccessful investment). Nevertheless, this prospect is dominated by the expected payoff from transferring the whole revenue to the predecessor (knowing that your potential successor will do the same). In that sense, it is for the common (greater) good that one individual is sacrificed.

The following result is an indirect consequence of Theorem 2.

Proposition 1 *The socially optimal equilibrium involves underinvestment.*

The intuition behind Proposition 1 is the following. When an agent makes an investment, she not only attracts the immediate successor, but also generates a positive externality so that her successor has the opportunity to further expand the hierarchy and generate more revenue. The first-best solution takes into account this positive externality and requests agents to make more investment compared with the decentralized case.

3.3 Decentralized initiator-optimality

The next result determines a one-parameter family of focal and simple initiator-optimal allocation schemes; namely, those imposing to each member a full transfer to two recipients (the predecessor and the initiator) with a constant ratio among the transfers. Formally, for each $\alpha \in [0, 1]$, the scheme t^α is constructed so that each agent (i) transfers a fraction α of the whole endowment (1) to the predecessor $(i-1)$, and the rest $(1-\alpha)$ to the initiator.[11] As Theorem 3

[11] Agent 1 transfers everything to the initiator. Note that, if $\alpha = 0$, the scheme would amount to transfer everything to the initiator, whereas if $\alpha = 1$, the scheme would amount to transfer everything to the immediate predecessor, which is precisely the scheme at Theorem 2.

states, for each technology, there exists a parameter so that the corresponding rule within the family is initiator-optimal.

Theorem 3 *There exists* $\alpha \in (0,1)$ *such that the allocation scheme* $t^{\alpha} = \{t^{\alpha}_{ij}\}_{\{ij\}\subset \mathbb{N}_0}$, *where*

$$t^{\alpha}_{ij} = \begin{cases} \alpha & \text{if } i = j+1 \geq 2, \\ 1-\alpha & \text{if } i \geq 2 \text{ and } j = 0, \\ 1 & \text{if } i \leq 1 \text{ and } j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

is initiator-optimal.

The proof of Theorem 3 (which also appears in the Appendix) is silent about the construction of the optimal ratio (α^*) describing the corresponding optimal scheme. The optimal α^* can be found as the solution to the problem

$$\max_{\alpha\in[0,1]} \Phi(\alpha),$$

where $\Phi : [0,1] \to \mathbb{R}$ is such that, for each $\alpha \in [0,1]$,

$$\Phi(\alpha) = (1 - x^{\alpha}_0) + p(x^{\alpha}_0)\left(\frac{1-\alpha p(x^{\alpha}_1)}{1-p(x^{\alpha}_1)}\right), \tag{4}$$

and x^{α}_1 is such that

$$p'(x^{\alpha}_1) = \frac{1}{\alpha}, \tag{5}$$

whereas x^{α}_0 is such that

$$p'(x^{\alpha}_0) = \frac{1-p(x^{\alpha}_1)}{1-\alpha p(x^{\alpha}_1)}. \tag{6}$$

Note that (5) guarantees that all followers are selecting their optimal investment, given the scheme t^{α}, whereas (6) guarantees that the initiator is selecting her optimal investment, given the scheme t^{α}.[12] Finally, (4) reflects the expected payoff for the initiator, at those investment choices, i.e., $\Phi(\alpha) = \mathbb{E}_0(t^{\alpha}, (x^{\alpha}_0, x^{\alpha}_1, x^{\alpha}_1, \dots))$.

Unlike the socially optimal scheme characterized in Theorem 2, an initiator-optimal scheme is not uniquely characterized. In fact, the important characteristic of an initiator-optimal scheme is that all agents, except for the initiator, have the same expected payoff in equilibrium (conditional on realization). For example, as shown in the next result, a member of the so-called *bubbling-up* (or geometric) schemes mentioned in the introduction, suitably reinterpreted

[12]It is not difficult to show that the extreme cases ($\alpha = 0$ and $\alpha = 1$) cannot yield the optimal scheme, which justifies the use of (5), the optimal condition for interior solutions.

from Hougaard et al., (2017), is also initiator-optimal.[13] Yet, the scheme defined in Theorem 3 remains intuitive and it is also simple to compute for given technologies. Another advantage, as shown in the proof of Theorem 3, is that all agents (except for the initiator) have a dominant strategy in the induced game.

Formally, for each $\lambda \in [0,1]$, the scheme t^λ is constructed so that each agent (i) transfers the whole endowment (1) to the predecessor $(i-1)$, who keeps a fraction λ and transfers the rest $(1-\lambda)$ to the predecessor, who keeps a fraction λ and transfers the rest $((1-\lambda)^2)$ to the predecessor, and so forth, with the initiator getting the residual. More precisely, the initiator obtains from each agent $i \geq 1$, the amount $(1-\lambda)^{i-1} = 1 - \lambda \sum_{k=0}^{i-2}(1-\lambda)^k$.[14]

We have the following result.

Proposition 2 *There exists $\lambda \in (0,1)$ such that the scheme $t^\lambda = \{t_{ij}^\lambda\}_{\{ij\} \subset \mathbb{N}_0}$, where*

$$t_{ij}^\lambda = \begin{cases} \lambda(1-\lambda)^{i-j-1} & \text{if } i > j > 0, \\ (1-\lambda)^{i-1} & \text{if } i > 1 \text{ and } j = 0, \\ 1 & i \leq 1 \text{ and } j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

is initiator-optimal.

Proposition 2 provides a formal verification of the claim made in Pikard et al., (2011) that bubbling-up schemes can provide optimal incentives for social mobilization and maximize the expected income of the initiator (i.e., the chance of winning in their case). We acknowledge that our setting is not equivalent to that in Pickard et al., (2011), but it is somewhat reminiscent. To wit, in Pickard et al., (2011), they work with *diffusion-based task environments* which consist of a set of agents, a set of edges characterizing social relationships between agents; a set of tasks; a technology function that returns the success probability of a given agent in executing a given task; and a budget that can be spent by the mechanism. In these environments, agents are not aware of the tasks a priori. Instead, they become aware of as a result of being directly informed by the mechanism through advertising; or being informed through recruitment by an

[13] In the original definition of a bubbling-up rule, each agent is allowed to keep a fraction λ of her value. Yet, in our model, such a scheme can never be optimal as letting agents keep fractions of their own value represents "dead capital" in terms of providing agents with incentives to invest. Thus, a natural reinterpretation is that each agent transfers the whole endowment to the predecessor, who keeps a fraction λ and *bubbles-up* the rest.

[14] Note that, when $\lambda = 0$, the scheme simply amounts to transfer everything to the initiator, whereas when $\lambda = 1$, the scheme simply amounts to transfer everything to the immediate predecessor.

acquaintance agent. Also, when a task is completed, the mechanism is able to identify not only the agent who executed it, but also the information pathway that led to that agent learning about the task. Thus, a mechanism in this setting specifies a set of initial nodes to target (e.g., via advertising) and the payment made to each agent, so that budget balance is guaranteed. Their *recursive incentive mechanism* divides the budget equally among all tasks and then allocates the amount corresponding to each task in a *geometric* way. That is, if B_i denotes the amount for task i, and agent j appears in position k in the sequence $\{1, 2, ..., n\}$, then j receives $\frac{B_i}{2^{n-k+1}}$ as payment. If we interpret that, in our setting, there are infinite sequences, each of them starting from the initiator, and going all the way down to each of the agents in the hierarchy, and that the budget for each of them is 1 (i.e., $B_i = 1$ for each $i = 1, 2, ...$), then the scheme formalized in Proposition 2, with $\lambda = \frac{1}{2}$, and the proviso that the budget surplus goes to the initiator, coincides with the *recursive incentive mechanism* from Pickard et al., (2011) just described. The popularity of this mechanism comes from its success in a real-life social mobilization experiment (the so-called DARPA Network Challenge). We illustrate here that, theoretically speaking, the mechanism can indeed be proven optimal in our setting (for the notion of initiator-optimality), but it is not unique in achieving that status.

4 Discussion

We have studied in this paper the optimal management of evolving hierarchies. While the allocation of profits among agents has long been studied in the literature, no theoretical studies exist dealing with the sharing of profits in order to maximize network expansions. Our stylized model captures necessary incentives in sharing rules for evolving priorities in order to maximize profits from a centralized and decentralized perspective. The canonical application of our model is the sharing of profits in multi-level marketing, which popular companies such as Avon, Herbalife, or Medlife use to entice sellers to expand the network of sellers.

In our model, an initiator invests (e.g., time or money) into finding a subordinate, who will bring revenues to the joint venture and who will invest herself into finding another subordinate, and so on. This process is stochastic and there is no amount of resources that guarantee that a successor will join. Furthermore, the probability of enrollment has decreasing marginal returns to investments. The initiator sets a transfer scheme specifying how revenues are reallocated, via upward transfers, as the hierarchy evolves. As mentioned above, our analysis is general and we do not need to specify a particular technology function to obtain our results.

In general, our model captures any type of evolving linear network process with autonomous nodes investing resources in continuing the hierarchical process. Our results demonstrate how reallocation of the revenues obtained by individual nodes influences investment decisions and welfare in the network. The model therefore has potentially important applications, relating to diverse areas such as network induction interventions, respondent-driven sampling, strategic recruiting in ongoing hierarchies, search and organizational hierarchy, incentive tree mechanisms, blockchain and Bitcoin mining, social mobilization, or resource allocation in the presence of a hierarchical structure. In particular, as mentioned above, our Theorem 2 can be viewed as formal support for the optimality of mining reward (*coinbase*) transactions in Bitcoin where the full reward goes solely to the successful miner. Furthermore, Proposition 2 can be seen a formal verification of the claim made in Pikard et al., (2011) that bubbling-up schemes can provide optimal incentives for social mobilization and maximize the expected income of the initiator.

It might be clarifying to mention that an equivalent interpretation of our model is that in which there are infinitely many agents formed in a hierarchical structure. Each agent makes effort to independently solve one task, and generates a certain reward if the task is successfully solved. The probability of success is increasing in the effort level. Notice that an agent is granted the chance to solve the task if and only if all of her predecessors successfully solve their own tasks. This alternative interpretation of the model may help to understand that in the social optimality problem and the initiator-optimality problem, (some) transfers go to the immediate predecessor. Winter (2006) explores a similar setting in which a project has to be managed by n individuals who act sequentially. Each individual has the option of investing at a cost. Investment will increase the probability that her task is successfully performed from $\delta \in [0, 1]$ to 1. For the project to succeed, all the tasks must succeed, in which case rewards are distributed to agents. He characterizes the optimal investment-inducing mechanism, i.e., the allocation of rewards that will induce all players to invest at minimal total cost to the principal, in which rewards increase with the cost and with δ (so that shirking becomes less attractive). More interestingly, late movers (whose investment decisions are observed by a small number of players) are given higher rewards than early movers (whose investment decisions are observed by many players) because the implicit threat against shirking is weaker.[15] Those considerations

[15] This implies that the working environment may be susceptible to incentive reversal, that is, increasing all agents' rewards may result in the shrinking of the set of agents who exert effort (see also Winter 2009). This is with a fixed structure of information about peers (either full information or no information at all). Winter (2010) compares between different information structures and studies their effect on agents' incentives.

cannot be brought to our setting with an infinite number of agents; thus, with no option to distinguish on the number of followers from one agent to another. Now, in Winter (2006), the equilibrium strategies in the optimal mechanism prescribe each agent to invest if and only if all agents she can observe invest as well. Thus, shirking by some players creates a domino effect, which induces all subsequent movers to shirk as well. This is indeed a somewhat similar feature to those exhibited by our optimal schemes in this setting.

We conclude mentioning that our analysis could be extended in several ways. We list here some natural ones.

First, we note that our model has concentrated on the linear-hierarchy case, and a natural next step is to extend the analysis to account for *branch hierarchies*, i.e., situations in which a given agent can have more than one immediate successor. For instance, one could think of recruiting teams rather than individuals, which might be necessary for some activities. In the benchmark case, all teams recruited in each step are of the same size (say, k), and each member of the team generates the same revenue ($1/k$ in step 1, $1/k^2$ in step 2, $1/k^3$ in step 3 and so on). This case could be easily addressed with the results from this paper.[16] The more general case in which teams recruited in each step can be of different size cannot be addressed in such a straightforward way and, thus, remains open for future research.

Second, as mentioned above, our model could be considered as a moral hazard optimal contracting problem, with the innovative aspect that the joint monetary outcome is an asymmetric function of the agents' (subordinates') efforts, due to the hierarchical structure. The case in which each subordinate writes her own contract (e.g., supply-chains) would be interesting too. Addressing this problem, which could be considered within the incipient area of mechanism games with multiple principals and agents (e.g., Yamashita, 2010), and also within the literature on hierarchical contracting (e.g., Mookherjee, 2006) is left for future research.

Finally, the investment level is private information in our model, and as such not viable to be contracted on. In order to bridge the gap between the first-best solution (with contractible effort) and the socially optimal solution (without contractible effort) we would need to consider intermediate scenarios regarding investment observability. One option would be to consider a public (observable) signal such as the overall investment from the initiator to the immediate predecessor. This is beyond the scope of this paper and, thus, it is also left for further research.

[16] This is essentially due to the fact that, in this case, there is a uniquely determined linear sub-hierarchy connecting each agent in the branch hierarchy to the initiator.

5 Appendix

Proof of Theorem 1

We prove first the existence of first-best socially optimal investment profiles. To do so, note first that, by letting x be a constant profile with sufficiently small investments, we can assume, without loss of generality, that the downward social values are non-negative, i.e., $\mathbb{V}^l(x_l, x_{l+1}, \dots) \geq 0$, for each period l.

Without loss of generality, we can restrict attention to profiles $x = (x_0, x_1, \dots)$ such that $x_i \leq 1$ for each $i = 0, \dots, \infty$. Otherwise, let i be the smallest integer for which $x_i > 1$. If $x_j > 1$ for all $j > i$, then obviously $\mathbb{V}(z^0) \geq \mathbb{V}(x)$, where z^0 is such that

$$z_k^0 = \begin{cases} x_k & \text{for each } k = 0, \dots, i-1, \\ 0 & \text{for each } k = i, i+1, \dots. \end{cases}$$

Alternatively, let j^1 be such that $x_{j^1} \leq 1 < x_j$ for each $j = i, \dots, j^1 - 1$, and let z^1 be such that

$$z_k^1 = \begin{cases} x_k & \text{for each } k = 0, \dots, i-1, \\ x_{j^1+k-i} & \text{for each } k = i, i+1, \dots. \end{cases}$$

Then, $\mathbb{V}(z^1) \geq \mathbb{V}(x)$. If $z_k^1 \leq 1$, for each $k > i$ (equivalently, $x_k \leq 1$, for each $k > j^1$) we have finished. Suppose, to the contrary, that there exists $k > i$ such that $z_k^1 > 1 \geq z_l^1$ for each $l = 1, \dots, k-1$. Then, we repeat the above process for z^1, instead of x, to construct a profile z^2 such that $z_l^2 \leq 1$ for each $l = 1, \dots, k$, and $\mathbb{V}(z^2) \geq \mathbb{V}(z^1) \geq \mathbb{V}(x)$. Eventually, this process will lead to a sequence of profiles, whose limit z^* is such that $z_k^* \leq 1$, for each $k = 0, \dots \infty$, and $\mathbb{V}(z^*) \geq \mathbb{V}(x)$.

Thus, the strategy space of each agent can be restricted to the interval $[0, 1]$. By Tychonoff's theorem, $[0, 1]^{\mathbb{N}}$ is a compact set. As $\mathbb{V}$ is a continuous real-valued function, when restricted to $[0, 1]^{\mathbb{N}}$, it follows (by the extreme value theorem), that $\mathbb{V}$ attains its least upper bound as its maximum, which proves existence.

Let $\bar{x}$ be a first-best optimal profile of investments. Clearly, $\bar{x}_i > 0$ for each $i \in \mathbb{N}_0$.[17] Suppose, by contradiction, that $\bar{x} = (\bar{x}_0, \bar{x}_1, \dots)$ is not constant. Then, there exists a pair $i, j \in \mathbb{N}_0$, such that $\bar{x}_i \neq \bar{x}_j$. As $\bar{x}$ is first-best socially optimal, it follows, by *dynamic consistency*, that

$$\mathbb{V}^{i+1}(\bar{x}_{i+1}, \bar{x}_{i+2}, \dots) = \mathbb{V}^{j+1}(\bar{x}_{j+1}, \bar{x}_{j+2}, \dots) = \mathbb{V}(\bar{x}),$$

[17] The assumption that $\lim_{x \to 0^+} p'(x_i) = +\infty$ implies that it is optimal to invest a positive amount.

which we denote by $\bar{\mathbb{V}}$. Note that $\bar{\mathbb{V}} > 0$, as $\bar{x}_i > 0$ for each $i \in \mathbb{N}_0$. This means, in particular, that $\bar{x}_i$ and $\bar{x}_j$ are the solutions for the problems

$$\arg\max p(x_l)\bar{\mathbb{V}} - x_l,$$

for $l = i, j$, whose first-order conditions are

$$p'(x_l)\bar{\mathbb{V}} - 1 = 0,$$

for $l = i, j$. Now, by the strict concavity of p, it follows that both solutions are the same, i.e., $\bar{x}_i = \bar{x}_j$, which is a contradiction.

In order to show uniqueness, note first that the function $f(\gamma) = \frac{1-\gamma}{1-p(\gamma)}$ has a unique maximum within the interval $[0, 1]$.[18] Let $\hat{\gamma}$ be such a maximum and let x be the corresponding constant investment profile. By dynamic consistency, a first-best optimal profile must be constant. By the above,

$$\mathbb{V}(x) = \frac{1-\hat{\gamma}}{1-p(\hat{\gamma})} > \frac{1-\gamma}{1-p(\gamma)} = \mathbb{V}(y),$$

for each $\gamma \in [0, 1] \setminus \{\hat{\gamma}\}$ and $y = (\gamma, \gamma, \dots)$. Thus, x is the unique first-best optimal profile, which concludes the proof of Theorem 1. □

Proof of Theorem 2

For each expected investment X, let $\mathbb{L}^{\max}(X)$ be the highest obtainable expected length from an investment profile x with expected investment $\mathbb{I}(x) = X$.

Step 0. $\mathbb{L}^{\max}$ *is well defined.*

For each investment profile x, and each $i \in \mathbb{N}$, let $a_i = \prod_{k=0}^{i-1} p(x_k)$, and $b_i = \prod_{k=0}^{i-1} p(x_k)x_i$. Note that,

$$\sum_{i=1}^{+\infty} a_i = \lim_{N\to\infty} \sum_{i=1}^{N} a_i = \lim_{N\to\infty} \sum_{i=1}^{N} \prod_{k=0}^{i-1} p(x_k) = \mathbb{L}(x) - 1,$$

and

$$\sum_{i=1}^{+\infty} b_i = \lim_{N\to\infty} \sum_{i=1}^{N} b_i = \lim_{N\to\infty} \sum_{i=1}^{N} \prod_{k=0}^{i-1} p(x_k)x_i = \mathbb{I}(x) - x_0.$$

It might well be the case that $\mathbb{L}(x) = \infty$ and $\mathbb{I}(x) = \infty$ (or, equivalently, that $\sum_{i=1}^{+\infty} a_i$ and $\sum_{i=1}^{+\infty} b_i$ are divergent series). Now, in order to show $\mathbb{L}^{max}$ is well defined, we prove that finite

[18] As $f(1) = 0$, $\lim_{\gamma\to 0^+} p'(\gamma) = +\infty$, and $p''(\gamma) < 0$, it follows that $p'(\gamma)f(\gamma) = 1$ (or, equivalently, $f'(\gamma) = 0$) has a unique solution.

expected investment implies finite expected length. This is equivalent to saying that if $\sum_{i=1}^{+\infty} b_i$ is a convergent series, then so is $\sum_{i=1}^{+\infty} a_i$. Assume then that $\sum_{i=1}^{+\infty} b_i$ is a convergent series.

Let N_1 denote the set of periods for which investment is lower (or equal) than 1, and N_2 denote the set of the remaining periods. That is, $N_1 = \{i \in \mathbb{N}_0 \text{ such that } x_i \leq 1\}$ and $N_2 = \mathbb{N}_0 \setminus N_1$. Let $N_1^i = N_1 \cap \{0, 1, \ldots, i-1\}$ and n_1^i be its cardinality. Then,

$$\sum_{i \in N_1} a_i = \sum_{i \in N_1} \prod_{k=0}^{i-1} p(x_k) < \sum_{i \in N_1} \prod_{k \in N_1^i} p(x_k) \leq \sum_{i \in N_1} p(1)^{n_1^i} \leq \sum_{i=1}^{+\infty} p(1)^i = \frac{1}{1-p(1)} < +\infty,$$

where the inequalities follow from the fact that $p(\gamma) < 1$ for each $\gamma \in \mathbb{R}_+$, and that $x_i \leq 1$ for each $i \in N_1^i$.

On the other hand, as $a_i < a_i x_i = b_i$, for each $i \in N_2$, and $\sum_{i=1}^{+\infty} b_i$ is a convergent series, it follows that so is $\sum_{i \in N_2} b_i$. Thus, by the comparison test, $\sum_{i \in N_2} a_i$ is also a convergent series. Thus, $\sum_{i=1}^{+\infty} a_i = \sum_{i \in N_1} a_i + \sum_{i \in N_2} a_i$, and it is also a convergent series, as desired.

We now show that $\mathbb{L}^{\max}(\mathbb{X})$ exists for any given X. For that, it is sufficient to show that each agent's investment can be restricted to a bounded interval, without loss of generality. To do so, we aim to show that, for each x with $\mathbb{I}(x) = X$, there exists z such that $z_t \leq X$ for each t, $\mathbb{I}(z) = X$, and $\mathbb{L}(z) \geq \mathbb{L}(x)$. Note that if $\mathbb{L}(x) \leq 1 + p(X)$, it suffices to consider $z = (X, 0, 0, \cdots)$, as $\mathbb{I}(z) = X = \mathbb{I}(x)$ and $\mathbb{L}(z) = 1 + p(X)$. Thus, in what follows, we suppose that $\mathbb{L}(x) > 1 + p(X)$. Or, equivalently, $\frac{\mathbb{L}(x)-1}{\mathbb{I}(x)} > \frac{p(X)}{X}$.[19]

Let $\mathbb{L}^t(x)$ denote the expected length from period t onwards, i.e.,

$$\mathbb{L}^t(x) = 1 + p(x_t) + p(x_t)p(x_{t+1}) + \cdots = 1 + \sum_{i=t}^{+\infty} \prod_{k=t}^{i} p(x_k).$$

And let $\mathbb{I}^t(x)$ denote the expected investment from period t onwards, i.e.,

$$\mathbb{I}^t(x) = x_t + p(x_t)x_{t+1} + p(x_t)p(x_{t+1})x_{t+2} + \cdots = x_t + \sum_{i=t}^{+\infty} \prod_{k=t}^{i} p(x_k)x_i.$$

The ratio $\frac{\mathbb{L}^t(x)-1}{\mathbb{I}^t(x)}$ represents the expected length obtained per expected investment unit, from period t onwards. Comparing this ratio with $\frac{p(X)}{X}$, we determine whether the *tails* at each period t are *efficient* or not.

Now, let x be an investment profile such that $\mathbb{I}(x) = X$, and for which there exists t such that $x_t > X \geq x_s$, for each $s \leq t$. Let x^η be the investment profile resulting from x after

[19] In what follows, we shall also make repeated use of two general properties of inequalities. More precisely, let $\varepsilon, \varsigma, \tau, \upsilon, \kappa \in \mathbb{R}_{++}$ such that $\varepsilon > \tau$, $\varsigma > \upsilon$, $\kappa > \frac{\varepsilon}{\varsigma}$ and $\kappa < \frac{\tau}{\upsilon}$. Then $\kappa > \frac{\varepsilon-\tau}{\varsigma-\upsilon}$ and $\frac{\tau}{\upsilon} > \frac{\varepsilon+\tau}{\varsigma+\upsilon}$.

increasing x_0, by η, where $0 < \eta \leq X - x_0$. As $\frac{\mathbb{L}(x)-1}{\mathbb{I}(x)} > \frac{p(X)}{X}$, it follows that

$$\frac{\partial \mathbb{L}(x^\eta)}{\partial \eta} \geq p'(X) \text{ , for all } 0 < \eta \leq X - x_0. \tag{7}$$

We now distinguish two cases:

Case 1. $\frac{\mathbb{L}^{t+1}(x)-1}{\mathbb{I}^{t+1}(x)} \leq \frac{p(X)}{X}$.

Let z be an investment profile such that

$$z_k = \begin{cases} x_0 + \varepsilon & \text{for } k = 0, \\ x_k & \text{for } k = 1, \ldots t, \\ 0 & \text{otherwise,} \end{cases}$$

where $\varepsilon > 0$ is such that $\mathbb{I}(z) = \mathbb{I}(x) = X$. Note that, by (7), because we cut off the inefficient tail and add instead to the investment at time zero, $\mathbb{L}(z) > \mathbb{L}(x)$. Now, as $z_t = x_t > X$, we have $\frac{\mathbb{L}^t(z)-1}{\mathbb{I}^t(z)} = \frac{p(x_t)}{x_t} \leq \frac{p(X)}{X}$. Thus, we can again create a new profile z' from z where $z_t = x_t$ is replaced with 0 and z_0 is increased until the expected investment is the same. Again, this new profile will have a greater length.

Case 2. $\frac{\mathbb{L}^{t+1}(x)-1}{\mathbb{I}^{t+1}(x)} \geq \frac{p(X)}{X}$.

In this case, we aim to show that the new profile obtained from "skipping" x_t, and then adjusting x_0 so that the expected investment is unchanged, has greater length. For this, note that

$$\mathbb{L}^t(x) = 1 + p(x_t)\mathbb{L}^{t+1}(x),$$

and

$$\mathbb{I}^t(x) = x_t + p(x_t)\mathbb{I}^{t+1}(x).$$

Thus,[20]

$$\frac{\mathbb{L}^t(x) - 1}{\mathbb{I}^t(x)} \leq \frac{\mathbb{L}^{t+1}(x) - 1}{\mathbb{I}^{t+1}(x)}.$$

Note also that

$$\mathbb{L}^t(x) - \mathbb{L}^{t+1}(x) = 1 - (1 - p(x_t))\mathbb{L}^{t+1}(x) \tag{8}$$

and

$$\mathbb{I}^t(x) - \mathbb{I}^{t+1}(x) = x_t - (1 - p(x_t))\mathbb{I}^{t+1}(x). \tag{9}$$

If $\mathbb{I}^t(x) \leq \mathbb{I}^{t+1}(x)$, then by the above observation that $\frac{\mathbb{L}^t(x)-1}{\mathbb{I}^t(x)} \leq \frac{\mathbb{L}^{t+1}(x)-1}{\mathbb{I}^{t+1}(x)}$, it follows that, by skipping x_t and possibly reducing investment in x_{t+1} until the same expected investment has been obtained, we can increase length.

[20] This holds provided $\frac{\mathbb{L}^{t+1}(x)-1}{\mathbb{I}^{t+1}(x)} \geq \frac{p(x_t)}{x_t}$, which follows because of the premise at Case 1, and the fact that we are assuming that x is a profile with $x_t > X$. Thus, by the strict concavity of p, $\frac{p(X)}{X} > \frac{p(x_t)}{x_t}$.

If $\mathbb{I}^t(x) > \mathbb{I}^{t+1}(x)$, and $\mathbb{L}^t(x) \leq \mathbb{L}^{t+1}(x)$, we can clearly increase the length by increasing investments (for agent 0, say) until the original level is obtained, and we are done.

Finally, if $\mathbb{I}^t(x) > \mathbb{I}^{t+1}(x)$, and $\mathbb{L}^t(x) > \mathbb{L}^{t+1}(x)$, by (8), (9), we obtain[21]

$$\frac{\mathbb{L}^t(x) - \mathbb{L}^{t+1}(x)}{\mathbb{I}^t(x) - \mathbb{I}^{t+1}(x)} = \frac{p(x_t) - (1 - p(x_t))(\mathbb{L}^{t+1}(x) - 1)}{x_t - (1 - p(x_t))\mathbb{I}^{t+1}(x)} < \frac{p(X)}{X} \tag{10}$$

This means that, by skipping x_t, one can increase the investment in x_0 until the same level of expected investment has been obtained, while the expected length has been increased.

The argument above can be replicated for eventual subsequent periods with investments above X.

Step 1. $\mathbb{L}^{\max}$ *is a strictly concave and differentiable function.*

Let x and y be investment profiles, with expected investments $\mathbb{I}(x) = X$ and $\mathbb{I}(y) = Y$, such that $\mathbb{L}(x) = \mathbb{L}^{\max}(X)$ and $\mathbb{L}(y) = \mathbb{L}^{\max}(Y)$. By the strict concavity of p, there exists a unique $z_0 < \frac{1}{2}x_0 + \frac{1}{2}y_0$ such that $p(z_0) = \frac{1}{2}p(x_0) + \frac{1}{2}p(y_0)$. Let $s_0 = \frac{p(x_0)}{2p(z_0)}$. Then, $0 < s_0 < 1$ and $1 - s_0 = \frac{p(y_0)}{2p(z_0)}$.

Now, replace the initiator's investment at both profiles by z_0 and consider the lottery $\bar{z}^0$ arising from combining both resulting profiles with probabilities s_0 and $1 - s_0$, respectively. Formally, let $\bar{z}^0 = [(z_0, x^{|1}; s_0), (z_0, y^{|1}; 1 - s_0)]$. Then,

$$\mathbb{I}(\bar{z}^0) = s_0\mathbb{I}((z_0, x^{|1})) + (1 - s_0)\mathbb{I}((z_0, y^{|1})) = \frac{\mathbb{I}(x) + \mathbb{I}(y)}{2} + z_0 - \frac{1}{2}x_0 - \frac{1}{2}y_0 < \frac{\mathbb{I}(x) + \mathbb{I}(y)}{2},$$

and

$$\mathbb{L}(\bar{z}^0) = s_0\mathbb{L}((z_0, x^{|1})) + (1 - s_0)\mathbb{L}((z_0, y^{|1})) = \frac{\mathbb{L}(x) + \mathbb{L}(y)}{2}.$$

Similarly, by the strict concavity of p, there exists a unique $z_1 < s_0x_1 + (1 - s_0)y_1$ such that $p(z_1) = s_0p(x_1) + (1 - s_0)p(y_1)$. Let $s_1 = \frac{s_0p(x_1)}{p(z_1)}$. Then, $0 < s_1 < 1$ and $1 - s_1 = \frac{(1-s_0)p(y_1)}{p(z_1)}$.

Let $\bar{z}^1 = [(z_0, z_1, x^{|2}; s_1), (z_0, z_1, y^{|2}; (1 - s_1))]$. Then,

$$\begin{aligned}
\mathbb{I}(\bar{z}^1) &= s_1\mathbb{I}((z_0, z_1, x^{|2})) + (1 - s_1)\mathbb{I}((z_0, z_1, y^{|2})) \\
&= z_0 + p(z_0)z_1 + \frac{s_1p(z_0)p(z_1)}{p(x_0)p(x_1)}\left(\mathbb{I}(x) - x_0 - p(x_0)x_1\right) + \frac{(1 - s_1)p(z_0)p(z_1)}{p(y_0)p(y_1)}\left(\mathbb{I}(y) - y_0 - p(y_0)y_1\right) \\
&= z_0 + p(z_0)z_1 + \frac{1}{2}\left(\mathbb{I}(x) - x_0 - p(x_0)x_1\right) + \frac{1}{2}\left(\mathbb{I}(y) - y_0 - p(y_0)y_1\right) \\
&= \frac{\mathbb{I}(x) + \mathbb{I}(y)}{2} + \left(z_0 - \frac{1}{2}x_0 - \frac{1}{2}y_0\right) + \left(p(z_0)z_1 - \frac{1}{2}p(x_0)x_1 - \frac{1}{2}p(y_0)y_1\right) \\
&< \frac{\mathbb{I}(x) + \mathbb{I}(y)}{2},
\end{aligned}$$

[21] Take, in Footnote 18, $\kappa = \frac{p(X)}{X}$, $\varepsilon = p(x_t)$, $\varsigma = x_t$, $\tau = (1 - p(x_t))(\mathbb{L}^{t+1}(x) - 1)$, and $\upsilon = (1 - p(x_t))\mathbb{I}^{t+1}(x)$. Then, obviously, $\kappa > \frac{\varepsilon}{\varsigma} = \frac{p(x_t)}{x_t}$, and $\kappa < \frac{\tau}{\upsilon} = \frac{\mathbb{L}^{t+1}(x) - 1}{\mathbb{I}^{t+1}(x)}$, provided we are assuming that x is a profile with $x_t > X$.

and

$$\begin{aligned}
\mathbb{L}(\bar{z}^1) &= s_1\mathbb{L}((z_0, z_1, x^{|2})) + (1-s_1)\mathbb{L}((z_0, z_1, y^{|2})) \\
&= 1 + p(z_0) + \frac{s_1 p(z_0)p(z_1)}{p(x_0)p(x_1)}\left(\mathbb{L}(x) - 1 - p(x_0)\right) + \frac{(1-s_1)p(z_0)p(z_1)}{p(y_0)p(y_1)}\left(\mathbb{L}(y) - 1 - p(y_0)\right) \\
&= 1 + p(z_0) + \frac{1}{2}\left(\mathbb{L}(x) - 1 - p(x_0)\right) + \frac{1}{2}\left(\mathbb{L}(y) - 1 - p(y_0)\right) \\
&= \frac{\mathbb{L}(x) + \mathbb{L}(y)}{2},
\end{aligned}$$

Let $\bar{z}$ denote the limit of the sequence $\bar{z}^0$, $\bar{z}^1$, $\bar{z}^2, \ldots$ constructed as above. Thus,

$$\mathbb{I}(\bar{z}) < \frac{\mathbb{I}(x) + \mathbb{I}(y)}{2},$$

and

$$\mathbb{L}(\bar{z}) = \frac{\mathbb{L}(x) + \mathbb{L}(y)}{2}.$$

Finally, let $\varepsilon > 0$ be such that

$$\bar{z}_0 + \varepsilon + \frac{p(\bar{z}_0 + \varepsilon)}{p(\bar{z}_0)}\left(\mathbb{I}(\bar{z}) - \bar{z}_0\right) = \frac{1}{2}\left(\mathbb{I}(x) + \mathbb{I}(y)\right),$$

and z^* be the profile such that

$$z_k^* = \begin{cases} \bar{z}_k + \varepsilon & \text{for } k = 0, \\ \bar{z}_k & \text{otherwise.} \end{cases}$$

Then,

$$\mathbb{I}(z^*) = \bar{z}_0 + \varepsilon + \frac{p(\bar{z}_0 + \varepsilon)}{p(\bar{z}_0)}\left(\mathbb{I}(\bar{z}) - \bar{z}_0\right) = \frac{\mathbb{I}(x) + \mathbb{I}(y)}{2},$$

and

$$\mathbb{L}(z^*) \geq \mathbb{L}(\bar{z}) = \frac{\mathbb{L}(x) + \mathbb{L}(y)}{2}.$$

The above shows that $\mathbb{L}^{\max}$ is strictly concave. We conclude this step showing that it is also differentiable. Fix $X > 0$. As $\mathbb{L}^{\max}$ is strictly concave, both the left and the right derivative at X exist, i.e., $\mathbb{L}^{\max}_{-}(X)$ and $\mathbb{L}^{\max}_{+}(X)$. Thus, it suffices to show that $\mathbb{L}^{\max}_{-}(X) = \mathbb{L}^{\max}_{+}(X)$. Let x be an investment profile such that $\mathbb{I}(x) = X$ and $\mathbb{L}(x) = \mathbb{L}^{\max}(X)$. As p is differentiable, $\mathbb{L}(x)$ is differentiable at x_0. As $\frac{\partial \mathbb{L}(x)}{\partial x_0} \leq \mathbb{L}^{\max}_{+}(X)$, and $\frac{\partial \mathbb{L}(x)}{\partial x_0} \geq \mathbb{L}^{\max}_{-}(X)$, it follows that $\mathbb{L}^{\max}_{+}(X) \geq \mathbb{L}^{\max}_{-}(X)$. As $\mathbb{L}^{\max}$ is strictly concave, $\mathbb{L}^{\max}_{+}(X) \leq \mathbb{L}^{\max}_{-}(X)$. Thus, altogether, $\mathbb{L}^{\max}_{-}(X) = \mathbb{L}^{\max}_{+}(X)$, as desired.

Step 2. For each non-constant investment profile x, there exists a (unique) constant profile y such that $\mathbb{L}(x) = \mathbb{L}(y)$ and $\mathbb{I}(y) < \mathbb{I}(x)$.

As mentioned at Step 0, $\mathbb{L}^l(x) = 1 + p(x_l)\mathbb{L}^{l+1}(x)$, for each investment profile x, and each period l. Thus, $\mathbb{L}^l(x) = \mathbb{L}^{l'}(x)$, for all periods l and l', if and only if x is a constant profile. Consequently, if x is a non-constant profile, there exists l such that $\mathbb{L}(x) \neq \mathbb{L}^l(x)$. Without loss of generality, assume $\mathbb{L}(x) < \mathbb{L}^l(x)$.[22] As the derivative of $\mathbb{L}^{\max}$ at $\mathbb{I}(x)$ is $\frac{\partial \mathbb{L}(x)}{\partial x_0}$, it follows, by the strict concavity of $\mathbb{L}^{\max}$, that it is possible to marginally decrease x_l and increasing x_0 while keeping the expected length constant and get a strictly lower expected investment. From this, it follows that x can also be *improved* upon constructing another profile with the same expected investment and strictly greater expected length. Thus if $\mathbb{L}(x) = \mathbb{L}^{\max}(\mathbb{I}(x))$, then x must be a constant investment profile.

Conversely, $\mathbb{L}(y) = \mathbb{L}^{\max}(\mathbb{I}(y))$ for each constant profile y. Otherwise, if another profile z with $\mathbb{I}(z) = \mathbb{I}(y)$ yields a maximal possible expected length greater that $\mathbb{L}(y)$ then z must also be a constant profile (contradicting that they have same expected investment).

To conclude, let x be an arbitrary non-constant profile. Let y be the (unique) constant profile satisfying that $\mathbb{L}(y) = \mathbb{L}(x)$. It follows from the above that $\mathbb{I}(y) < \mathbb{I}(x)$, as desired.

Step 3. For each non-constant equilibrium profile x, there exists a (unique) constant equilibrium profile y such that $\mathbb{L}(x) = \mathbb{L}(y)$ and $\mathbb{V}(x) < \mathbb{V}(y)$.

Let x be a non-constant equilibrium profile, for a certain transfer scheme t. Without loss of generality, suppose that $t_{ii} = 0$ for each $i \in \mathbb{N}$ (recall that $t_{00} = 1$, by definition). For each $i \in \mathbb{N}_0$, let $\mathbb{E}_i^*(t, x)$ denote i's expected gross return from investment conditional on realization of the next agent $(i + 1)$. Formally,

$$\mathbb{E}_i^*(t, x) = \frac{1}{p(x_i)} \sum_{l=1}^{+\infty} \prod_{k=0}^{l-1} p(x_{i+k}) t_{(i+l)i}.$$

Note that

$$p(x_0)\mathbb{E}_0^*(t, x) + p(x_0)p(x_1)\mathbb{E}_1^*(t, x) + \cdots = \mathbb{L}(x) - 1.$$

Furthermore, for each $i \in \mathbb{N}_0$, let $e(x_i)$ denote the amount of expected transfer to agent i, conditional upon realization of agent $i + 1$ which incentivizes investment x_i, i.e., $e(x_i) = \frac{1}{p'(x_i)}$ and $e(0) = 0$. Let $z = (\xi, \xi, \dots)$ be the constant profile such that

$$p(\xi)e(\xi) + p(\xi)p(\xi)e(\xi) + \cdots = p(x_0)\mathbb{E}_0^*(t, x) + p(x_0)p(x_1)\mathbb{E}_1^*(t, x) + \dots \tag{11}$$

Thus,

$$e(\xi) = \frac{\mathbb{L}(x) - 1}{\mathbb{L}(z) - 1}.$$

[22] The case $\mathbb{L}(x) > \mathbb{L}^l(x)$ is similar.

We distinguish two cases.

Case 1: $\mathbb{I}(z) \geq \mathbb{I}(x)$.

Let z' be a constant profile such that $\mathbb{I}(z') = \mathbb{I}(x)$. It follows, by Step 2, that $\mathbb{L}(z') \geq \mathbb{L}(x)$. Hence, $\mathbb{L}(z) \geq \mathbb{L}(x)$. Thus, $e(\xi) \leq 1$. The profile z can therefore be supported by the scheme that transfers $e(\xi)$ from agent $l+1$ to l, provided agent $l+1$ is realized. As $\mathbb{I}(z') \leq \mathbb{I}(z)$, it follows that z' can also be supported. By continuity of $\mathbb{L}^{\max}$ it follows that there exists a constant profile y such that $\mathbb{L}(y) = \mathbb{L}(x)$ and $\mathbb{I}(y) < \mathbb{I}(z') = \mathbb{I}(x)$. As z' can be supported in equilibrium, so can y.

Case 2: $\mathbb{I}(z) < \mathbb{I}(x)$.

If $\mathbb{L}(z) \geq \mathbb{L}(x)$ then z can be supported as described in Case 1 and the proof of Step 3 would equivalently be completed. Thus, we assume that $\mathbb{L}(z) < \mathbb{L}(x)$, i.e., $e(\xi) > 1$. Let z' be the constant profile such that $\mathbb{L}(z') = \mathbb{L}(x) > \mathbb{L}(z)$. Then, $z'_0 = z'_1 = z'_2 = \cdots = \xi' > \xi$ and, therefore, $e(\xi') > e(\xi) > 1$.

Let $\omega_0 = 1$, $\omega_1 = p(x_0)$, $\omega_2 = p(x_0)p(x_1)$, etc. Then, $\sum_{k=0}^{\infty} \omega_k = \mathbb{L}(x) = \mathbb{L}(z')$, and $\sum_{k=0}^{\infty} \omega_k p(x_k) = \mathbb{L}(x) - 1 = \mathbb{L}(z') - 1 = p(\xi')\mathbb{L}(z')$, which implies that $p(\xi') = \frac{\sum_{k=0}^{\infty} \omega_k p(x_k)}{\sum_{k=0}^{\infty} \omega_k}$. As p is strictly concave, it follows that $\xi' < \frac{\sum_{k=0}^{\infty} \omega_k x_k}{\sum_{k=0}^{\infty} \omega_k}$. Furthermore, as $g(\gamma) = \frac{p(\gamma)}{p'(\gamma)} = p(\gamma)e(\gamma)$ is a strictly increasing and convex function, it follows that

$$g(\xi') < g\left(\frac{\sum_{k=0}^{\infty} \omega_k x_k}{\sum_{k=0}^{\infty} \omega_k}\right) < \frac{\sum_{k=0}^{\infty} \omega_k g(x_k)}{\sum_{k=0}^{\infty} \omega_k}.$$

Thus,

$$\sum_{k=0}^{\infty} \omega_k p(\xi')e(\xi') < \sum_{k=0}^{\infty} \omega_k p(x_k)e(x_k) = \mathbb{L}(x) - 1,$$

which implies that $e(\xi') < 1$, a contradiction.

Step 4. The allocation scheme in the statement is socially optimal.

Note first that a socially optimal allocation scheme x^* involves null self-transfers. By Step 3, a socially optimal equilibrium must be constant. Let ξ be the (constant) investment associated with the socially optimal equilibrium. It follows from Step 3 that $e(\xi) \leq 1$. Then, the scheme that transfers $e(\xi)$ from agent $l+1$ to l, provided agent $l+1$ is realized, is well-defined and supports that equilibrium. Note that, as the equilibrium of a constant scheme involves that $\mathbb{E}_i^*$ is constant across agents, it follows that $\mathbb{E}_i^* = 1$, for each i, as this is the only way that the identity $p(\xi)\mathbb{E}_0^* + p(\xi)p(\xi)\mathbb{E}_1^* + \cdots = \mathbb{L}(x^*) - 1$ can hold. Altogether, it shows that t^* is the socially optimal scheme.

Step 5. Uniqueness.

Let x^* denote the decentralized socially optimal (constant) equilibrium profile induced by t^*. By contradiction, assume that there exists another decentralized socially optimal scheme t with equilibrium profile x^*. As agent 1 can only transfer value to agent 0, or herself, we can increase social value by increasing t_{10} until we reach $t_{10} = 1$. Moreover, $t_{j0} = 0$, for each $j \geq 2$ (as, otherwise, agent 0 would over-invest). By induction, we obtain that $t = t^*$, a contradiction. □

Proof of Proposition 1

Let $z^* = (\xi, \xi, \dots)$ be the (constant) first-best socially optimal investment profile and $x^* = (\chi, \chi, \dots)$ be the socially optimal equilibrium investment profile (which is also constant, by Theorem 2). As shown in the proof of Theorem 1, $p'(\xi) = \frac{1-p(\xi)}{1-\xi}$ and $\frac{1-\xi}{1-p(\xi)} = \mathbb{V}(z^*) > \mathbb{V}(0) = 1$. Thus, $\xi < p(\xi)$. Now, by Theorem 2, $p'(\chi) = 1$. Thus, $p'(\xi) = \frac{1-p(\xi)}{1-\xi} < 1 = p'(\chi)$. Therefore, by strict concavity of p, $\xi > \chi$, as desired. □

Proof of Theorem 3

Step 1. There exists an initiator-optimal allocation scheme.

First we show that, without loss of generality, investments are bounded. For this, recall that we know from Theorem 2 that the socially optimal equilibrium investments x are constant over time. In particular, $\mathbb{L}(x)$ is finite.

As in the proof of Theorem 2, let $\mathbb{E}^*_i(t, x)$ denote i's expected gross return from investment conditional on realization of the next agent. Formally, $\mathbb{E}^*_i(t, x) = \frac{1}{p(x_i)} \sum_{l=1}^{+\infty} \prod_{k=0}^{l-1} p(x_{i+k}) t_{(i+l)i} = \sum_{l=1}^{+\infty} \prod_{k=1}^{l-1} p(x_{i+k}) t_{(i+l)i}$. Furthermore, let $e(\mathbb{E}^*)$ denote the optimal investment for an agent, provided that the agent receives an expected payoff of $\mathbb{E}^*$ if successful in obtaining the next agent.

We claim that, without loss of generality, we can restrict attention to a bounded set of potential investments. In particular, for each agent we can restrict attention to investments within $[0, e(\mathbb{L}(x))]$, where $\mathbb{L}(x)$ is the expected length of the hierarchy under socially optimal equilibrium investments x.

By contradiction, suppose that, for an initiator-optimal scheme t, an equilibrium y exists such that $y_i > e(\mathbb{L}(x))$. As y is initiator-optimal, $y_0 \geq y_i > e(\mathbb{L}(x))$.[23] Now, consider the subgame starting from agent 1, and the scheme obtained from t by redirecting any transfer to

[23] Otherwise, the whole scheme could be truncated, assigning the residual transfers to i.

agent 0 to agent 1 instead. The social value of the equilibrium of the subgame starting from agent 1 (obtained after appropriately increasing the investment of agent 1 until equilibrium is obtained) is higher than the social value of x.[24] This contradicts that x is a socially optimal equilibrium.

Thus, without loss of generality, we can restrict attention to investments between 0 and $\mathbb{L}(x)$ for all agents.

Next, observe that the set of equilibria (and associated schemes) is a closed set.[25] As it is also a bounded set, it follows that there exists an equilibrium x^* with an associated scheme t^* such that $\mathbb{E}_0(t^*, x^*) \geq \mathbb{E}_0(t, x)$, for each equilibrium x in the game induced by scheme t.

Step 2. There exists $\alpha \in (0,1)$ *such that the scheme defined by*

$$t_{ij}^{\alpha} = \begin{cases} \alpha & \text{if } i = j+1 \geq 2, \\ 1-\alpha & \text{if } i \geq 2 \text{ and } j = 0, \\ 1 & \text{if } i \leq 1 \text{ and } j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

is initiator-optimal

Let t be an initiator-optimal allocation scheme. We now claim some properties of t.

First, $t_{ii} = 0$, for each $i \in \mathbb{N}_0 \setminus \{0\}$. The reason is simply that a self-transfer is irrelevant for the investment decision. More precisely, if $t_{ii} > 0$, for some $i \in \mathbb{N}_0 \setminus \{0\}$, and x is an equilibrium of the induced game, then we could construct a transfer t', such that $t'_{ii} = 0$, $t'_{i0} = t_{i0} + t_{ii}$, and $t'_{jk} = t_{jk}$ otherwise, and whose induced game would have x' as an equilibrium, such that $\mathbb{E}_0^*(t, x) < \mathbb{E}_0^*(t', x')$.

Let N_i^+ denote the set of i's followers (including i) in the hierarchy, i.e., $N_i^+ = \{i, i+1, \dots\}$. Let x be an initiator-optimal equilibrium in the game induced by t, and let $\mathbb{S}_i(t, x)$ denote the expected total transfers from members of coalition N_i^+ to the predecessors of i, given agent i

[24]Without $\mathbb{L}(x)$ to agent 1, the social value of the subgame of y starting from agent 1 is non-negative, and adding $\mathbb{L}(x)$ to agent 1 would increase the social value of the subgame with at least $\mathbb{L}(x)$, even before the increase in investment of agent 1.

[25]This is due to the fact that each agent's optimal investment is continuous in the transfers, and the Nash equilibrium inequality condition holds in all limit points.

joins the hierarchy. Formally,[26]

$$\begin{aligned}\mathbb{S}_i(t,x) &= (t_{i0}+t_{i1}+\cdots+t_{ii-1})+\\ &+ p(x_i)(t_{(i+1)0}+t_{(i+1)1}+\cdots+t_{(i+1)(i-1)})+\\ &+ p(x_i)p(x_{i+1})(t_{(i+2)0}+t_{(i+2)1}+\cdots+t_{(i+2)(i-1)})+\\ &\dots\\ &= \sum_{j=i}^{+\infty}\prod_{l=i}^{j-1}p(x_l)\sum_{k=0}^{i-1}t_{jk}.\end{aligned}$$

Let j be such that $0<j<i$. We claim that $\mathbb{S}_i(t,x)=\mathbb{S}_j(t,x)$.

Suppose, by contradiction, that $\mathbb{S}_i(t,x)>\mathbb{S}_j(t,x)$. Then, a new allocation scheme t' can be defined as follows. For each pair $k,l\in N_j^+$, with $l\le k$, let $t'_{kl}=t_{(k+(i-j))(l+(i-j))}$. As for the rest of the transfers, the only constraint is that $\mathbb{E}^*_m(t',x')=\mathbb{E}^*_m(t,x)$, for each $m=1,\dots,j-1$, and $\mathbb{E}^*_0(t',x')>\mathbb{E}^*_0(t,x)$. Note that such an allocation scheme t' is feasible due to the hypothesis that $\mathbb{S}_i(t,x)>\mathbb{S}_j(t,x)$.

As x was an initiator-optimal equilibrium of the game induced by t, the allocation scheme t' induces an equilibrium x' with the following features:

- $x'_0>x_0$ and $\mathbb{E}^*_0(t',x')>\mathbb{E}^*_0(t,x)$,
- $x'_l=x_l$ for each $l=1,2,\dots,j-1$,
- $x'_l=x_{i+j-l}$ for each $l=j,j+1,\dots$.

But then, because of the first item, t would not be initiator-optimal, a contradiction.

Suppose now that $\mathbb{S}_i(t,x)<\mathbb{S}_j(t,x)$. Then, a new allocation scheme t' can be defined as follows. For each pair $k,l\in N_i^+$, with $l\le k$, let $t'_{kl}=t_{(k+(j-i))(l+(j-i))}$. As for the rest of the transfers, the only constraint is that $\mathbb{E}^*_m(t',x')=\mathbb{E}^*_m(t,x)$, for each $m=1,\dots,i-1$, and $\mathbb{E}^*_0(t',x')>\mathbb{E}^*_0(t,x)$. Note that such an allocation scheme t' is feasible due to the hypothesis that $\mathbb{S}_i(t,x)<\mathbb{S}_j(t,x)$.

As x was an initiator-equilibrium of the game induced by t, the allocation scheme t' induces an equilibrium x' with the following features:

- $x'_0>x_0$ and $\mathbb{E}^*_0(t',x')>\mathbb{E}^*_0(t,x)$,
- $x'_l=x_l$ for each $l=1,2,\dots,i-1$,

[26] We consider the notational convention that $\prod_{l=j}^{j-1}p(x_l)=1$.

- $x'_l = x_{l-i+j}$ for each $l = i, i+1, \ldots$.

But then, because of the first item, t would not be initiator-optimal, a contradiction.

We therefore conclude that $\mathbb{S}_i(t,x) = \mathbb{S}_j(t,x)$.

Next, we argue that an initiator-optimal allocation scheme induces a symmetric equilibrium among the agents following the initiator. For this, suppose that t is initiator-optimal and x is a corresponding initiator-optimal equilibrium.

Now, suppose that for some $i \geq 1$ we have $x_i \neq x_{i+1}$.

We then argue that it is possible to change t to another allocation scheme $\hat{t}$ such that a new equilibrium $\hat{x}$ is obtained with $\hat{x}_i = \hat{x}_{i+1}$ and $\hat{x}_j = x_j$ for all $j \leq i$, and, in particular, the payoff to agent 0 is the same at $\hat{x}$ and at x.

For this, we establish $\hat{t}$ such that agent $i+1$ faces the same transfers from subsequent agents as i did under scheme t. More precisely, let $\hat{t}$ satisfy the following:

- $\hat{t}_{kj} = t_{(k-1)(j-1)}$, for each $k \geq j \geq i+1$,
- $\hat{t}_{kj} = t_{kj}$, for each $j \leq k \leq i+1$,

Now, let $\hat{x}$ be defined by $\hat{x}_j = x_j$ for $j \leq i$ and $\hat{x}_j = x_{j-1}$ for $j > i$. Then, as $\mathbb{S}_{i+1}(\hat{t},\hat{x}) = \mathbb{S}_i(t,x)$ it is possible to choose the remaining transfers $\hat{t}_{kj}$ (for $j \leq i$, $k \geq i+2$) such that each agent $j \leq i$ gets the same expected gross return from investment conditional on realization of the next agent with $\hat{x}_j$ under $\hat{t}$, as with x_j under t. This implies that $\hat{x}$ is an equilibrium under $\hat{t}$, from where we can conclude. This process can be sequentially repeated for increasing i, starting at $i = 1$. Thus, without loss of generality, we can assume that $x_i = x_j$ for each pair $i, j > 0$. Moreover, as (x,t) is initiator-optimal, $x_0 \geq x_i$ for each $i > 0$.

Now, for each $i > 0$, let $\alpha = \mathbb{E}_i^*(t,x)$. As $x_0 \geq x_i$ for each $i > 0$, it follows by Theorem 2, that $\alpha \leq 1$. It is straightforward to show that $\alpha = 1$ and $\alpha = 0$ cannot be optimal. Thus, $\alpha \in (0,1)$.

To conclude, consider the scheme $t^\alpha = \{t^\alpha_{ij}\}_{\{ij\}\subset\mathbb{N}_0}$, where

$$t^\alpha_{ij} = \begin{cases} \alpha & \text{if } i = j+1 \geq 2, \\ 1-\alpha & \text{if } i \geq 2 \text{ and } j = 0, \\ 1 & \text{if } i \leq 1 \text{ and } j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

This scheme induces a game with an equilibrium in which each agent (except for the initiator) invests x_i and gets an expected gross return from investment conditional on realization of the next agent of α. Thus, $\mathbb{E}_i(t^\alpha, x) = \mathbb{E}_i(t, x)$ for each $i > 0$. The initiator gets then the residual. As $\mathbb{E}_i^*(t^\alpha, x) = \mathbb{E}_i^*(t, x)$ for each $i > 0$, and $p(x_0)\mathbb{E}_0^*(t, x) + p(x_0)p(x_1)\mathbb{E}_1^*(t, x) + \cdots = \mathbb{L}(x) - 1$, it follows that $\mathbb{E}_0^*(t^\alpha, x) = \mathbb{E}_0^*(t, x)$. □

Proof of Proposition 2

Let us start considering the initiator-optimal scheme t^α from Theorem 3. The game induced by such a scheme involves that all agents (except for the initiator) invest x_1^α, the unique solution of the equation $p'(\gamma) = 1/\alpha$. As for the initiator, the optimal investment in that game is x_0^α, the solution to $p'(\gamma) = \frac{1-p(x_1^\alpha)}{1-\alpha p(x_1^\alpha)}$. Then, for each $\alpha \in (0, 1)$, the expected payoff for the initiator is given by

$$\mathbb{E}_0(t^\alpha, (x_0^\alpha, x_1^\alpha, x_1^\alpha, \dots)) \quad = \quad (1 - x_0^\alpha) + p(x_0^\alpha)\left(\frac{1 - \alpha p(x_1^\alpha)}{1 - p(x_1^\alpha)}\right)$$

Now, the expected return from such an investment for the initiator with the scheme t^λ is

$$\begin{aligned}
\mathbb{E}_0(t^\lambda, (x_0^\alpha, x_1^\alpha, x_1^\alpha, \dots)) &= (1 - x_0^\alpha) + p(x_0^\alpha) + p(x_0^\alpha)p(x_1^\alpha)(1-\lambda) + p(x_0^\alpha)p(x_1^\alpha)p(x_1^\alpha)(1-\lambda)^2 + \cdots \\
&= (1 - x_0^\alpha) + p(x_0^\alpha)\sum_{i=1}^{+\infty}\left((1-\lambda)p(x_1^\alpha)\right)^i. \\
&= (1 - x_0^\alpha) + \frac{p(x_0^\alpha)}{1 - (1-\lambda)p(x_1^\alpha)}.
\end{aligned}$$

Thus, it is straightforward to show that

$$\mathbb{E}_0(t^\alpha, (x_0^\alpha, x_1^\alpha, x_1^\alpha, \dots)) = \mathbb{E}_0(t^\lambda, (x_0^\alpha, x_1^\alpha, x_1^\alpha, \dots)),$$

if and only if

$$\frac{1 - \alpha p(x_1^\alpha)}{1 - p(x_1^\alpha)} = \frac{1}{1 - (1-\lambda)p(x_1^\alpha)}. \tag{12}$$

As the right hand side takes the value $\frac{1}{1-p(x_1^\alpha)}$, when $\lambda = 0$, and 1, when $\lambda = 1$, it follows that there exists a unique $\hat{\lambda} \in (0, 1)$ for which (12) holds. By (12), it also follows that, for each other agent i,

$$\mathbb{E}_i(t^{\hat{\lambda}}, x_1^\alpha) = p(x_1^\alpha)\hat{\lambda}\sum_{j=0}^{+\infty}((p(x_1^\alpha)(1-\hat{\lambda}))^j - x_1^\alpha = \frac{p(x_1^\alpha)\hat{\lambda}}{1 - p(x_1^\alpha)(1-\hat{\lambda})} - x_1^\alpha = \alpha p(x_1^\alpha) - x_1^\alpha = \mathbb{E}_i(t^\alpha, x_1^\alpha).$$

Altogether, this shows that, with a suitable choice of λ, the corresponding bubbling-up scheme can be initiator-optimal as well. □

References

[1] Aral, S., Walker, D., (2011) Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks. *Management Science* 57, 1623-1639.

[2] Azevedo E., Pennock, D., Waggoner, B., Weyl, G. (2020) Channel Auctions. *Management Science* 66, 2075-2082.

[3] Azoulay, P., (2004) Capturing Knowledge within and across Firm Boundaries: Evidence from Clinical Development. *American Economic Review* 94, 1591-1612.

[4] Bimpikis, K., Ehsani, S., Ilkilic, R., (2019) Cournot Competition in Networked Markets. *Management Science* 65, 2467–2481.

[5] Bird, C., (1976) On cost allocation for a spanning tree: a game theoretic approach. *Networks* 6, 335-350.

[6] Brocas, I., Carrillo, J., (2008) The Brain as a Hierarchical Organization. *American Economic Review* 98, 1312-1346.

[7] Catalini, C., Gans, J., (2020) Some simple economics of the blockchain. *Communications of the ACM* 63, 80-90.

[8] Chen, X., Papadimitriou, C., Roughgarden, T., (2019) An axiomatic approach to block rewards. *Proceedings of the 1st ACM Conference on Advances in Financial Technologies*, 124-131

[9] Cheng, S., De Franco, G., Jiang, H., Lin, P., (2019) Riding the Blockchain Mania: Public Firms' Speculative 8-K Disclosures. *Management Science* 65, 5901-5913.

[10] Claus, A., Kleitman, D., (1973) Cost allocation for a spanning tree. *Networks* 3, 289-304.

[11] Corominas-Murtra, B., Gonid, J., Sole R.V., Rodriguez-Caso, C., (2013) On the origins of hierarchy in complex networks. *Proceedings of the National Academy of Sciences* 110, 13316-13321.

[12] Dong, B., Ni, D., Wang, Y., (2012) Sharing a polluted river network. *Environmental and Resource Economics* 53, 367-387.

[13] Emek, Y., Karidi, R., Tennenholtz, M., Zohar, A., (2011) Mechanisms for multi-level marketing. *Proc. of 12th ACM Conference on Electronic Commerce*, 209-218.

[14] Furth, B. (ed) (2010) Handbook of Social Network Technologies and Applications, Springer.

[15] Galeotti, A., Golub, B., Goyal, S., Talamás, E., Tamuz, O., (2018) Optimal Taxation in Supply Chains. Mimeo. Harvard University.

[16] Goldlücke, S., (2017) Strategic recruiting in ongoing hierarchies. *Economics Letters* 156, 176-178.

[17] Harker, T., Vargas, L., (1987) The Theory of Ratio Scale Estimation: Saaty's Analytic Hierarchy Process. *Management Science* 33, 1383-1403.

[18] Heckathorn D., (1997) Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. *Social Problems* 44, 174-199.

[19] Holderness, C., Pontiff J., (2012) Hierarchies and the Survival of Prisoners of War During World War II. *Management Science* 58, 1873-1886.

[20] Holmstrom, B., (1982) Moral hazard in teams, *Bell Journal of Economics* 13, 324-340.

[21] Hougaard, J.L., (2018) Allocation in Networks. *MIT Press*. Cambridge, MA.

[22] Hougaard, J.L., Moreno-Ternero, J.D., Tvede, M., Østerdal, L.P., (2017) Sharing the proceeds from a hierarchical venture. *Games and Economic Behavior* 102, 98-110.

[23] Huberman, G., Leshno, J., Moallemi, C., (2019) An Economist's Perspective on the Bitcoin Payment System. *AEA Papers and Proceedings* 109, 93-96.

[24] Juarez, R., Ko, C-Y., Xue, J., (2018) Sharing sequential values in a network. *Journal of Economic Theory* 177, 734-779.

[25] Juarez, R., Nitta, K., Vargas, M., (2020) Profit-sharing and efficient time allocation. *Economic Theory* 70, 817-846.

[26] Littlechild, S., Owen. G., (1973) A simple expression for the Shapley value in a special case. *Management Science* 20, 370-372.

[27] Lv, Y., Moscibroda, T., (2013) Fair and resilient incentive tree mechanisms. *Proc. of the 2013 ACM symposium on Principles of distributed computing*, 230-239.

[28] Mesarovic, M.D., Macko, D., Takahara, Y., (1970) Theory of Hierarchical, Multilevel, Systems. *Academic Press.* New York.

[29] Mihm, J., Loch, C., Wilkinson, D., Huberman, B., (2010) Hierarchical Structure and Search in Complex Organizations. *Management Science* 56, 831–848

[30] Mookherjee, D., (2006) Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective. *Journal of Economic Literature* 44, 367-390.

[31] Moreno-Ternero J.D., Roemer J., (2006). Impartiality, priority, and solidarity and in the theory of justice. *Econometrica* 74, 1419-1427.

[32] Nakamoto, S., (2008) Bitcoin: A peer-to-peer electronic cash system, White paper.

[33] Ni, D., Wang, Y., (2007) Sharing a polluted river. *Games and Economics Behavior*, 60, 176-186.

[34] Pickard, G., Pan, W., Rahwan, I., Cebrian, M., Crane, R., Madan, A., Pentland, A., (2011) Time critical social mobilization. *Science* 334, 509-512.

[35] Rajasethupathy P., Sankaran S., Marshel J., Kim C., Ferenczi E., Lee S., Berndt A., Ramakrishnan C., Jaffe A., Liston C., Deisseroth K. (2015) Projections from neocortex mediate top-down control of memory retrieval. *Nature* 526, 653-659.

[36] Rivkin, J. W. (2000) Imitation of complex strategies. *Management Science* 46, 824-844.

[37] Saaty, T., (1986) Axiomatic Foundation of the Analytic Hierarchy Process. *Management Science* 32, 841-855.

[38] Salthe, S.N., (1985) Evolving hierarchical systems. *Columbia University Press.*

[39] Sapolsky, R., (2005) The Influence of Social Hierarchy on Primate Health. *Science* 308, 648-652.

[40] Schrijvers, O., Boneh, D., Roughgarden, T., (2017) Incentive Compatibility of Bitcoin Mining Pool Reward Functions. In: Grossklags J., Preneel B. (eds) Financial Cryptography and Data Security. FC 2016. Lecture Notes in Computer Science, vol 9603. Springer, Berlin, Heidelberg.

[41] Strausz, R., (1999) Efficiency in Sequential Partnerships. *Journal of Economic Theory* 85, 140-156.

[42] Topkis, (1979) Equilibrium Points in Non-Zero Sum n-Person Submodular Games. *SIAM J. of Control and Optimization*, 17, 773-787.

[43] Valente, T.W., (2012) Network Interventions. *Science* 337, 49-53.

[44] Winter, E., (2004) Incentives and discrimination, *American Economic Review* 94, 764-773.

[45] Winter, E., (2006) Optimal incentives for sequential production processes, *The RAND Journal of Economics* 37, 376-390.

[46] Winter, E., (2009) Incentive Reversal. *American Economic Journal: Microeconomics* 1, 133-147.

[47] Winter, E., (2010) Transparency and incentives among peers, *The RAND Journal of Economics* 41, 504–523

[48] Yamashita, T., (2010) Mechanism games with multiple principals and three or more agents, *Econometrica* 78, 791-801