Theoretical Economics 18 (2023), 633–668 1555-7561/20230633

Multilevel marketing: Pyramid-shaped schemes or exploitative scams?

Yair Antler
Coller School of Management, Tel Aviv University

Motivated by the growing discussion on the resemblance of multilevel marketing schemes to pyramid scams, we compare the two phenomena based on their underlying compensation structures. We show that a company can design a pyramid scam to exploit a network of agents with coarse beliefs and that this requires the company to charge the participants a license fee and pay them a recruitment commission for each of the people that they recruit and that their recruits recruit. We characterize the schemes that maximize a company's profit when it faces fully rational agents, and establish that the company never finds it profitable to charge them a license fee or pay them recruitment commissions.

Keywords. Multilevel marketing, pyramid schemes, speculative trade, misspecified models, analogy-based expectations.

JEL classification. D84, D85, D86.

1. Introduction

What delineates pyramid scams from legitimate multilevel marketing enterprises? Recent growth in the multilevel marketing (MLM) industry has raised the urgency of this question for consumer protection agencies.[1] MLM companies such as Avon, Amway, Herbalife, and Tupperware use independent representatives to sell their products to friends and acquaintances. They all promote the opportunity of starting one's own business and making extra income; however, some (e.g., Bort (2016)) view these companies as pyramid scams whose main purpose is to exploit vulnerable individuals.

The MLM industry's questionable legitimacy received considerable media attention following a recent FTC investigation against Herbalife (FTC (2016)). Identifying whether

Yair Antler: yair.an@gmail.com
I am grateful to Rani Spiegler for many insightful discussions and very detailed feedback with many fruitful suggestions. I thank Vellore Arthi, Benjamin Bachi, Daniel Bird, Kfir Eliaz, Erik Eyster, Daniel Fershtman, Michael D. Grubb, Paul Heidhues, Philippe Jehiel, Alessandro Lizzeri, Francesco Nava, Ronny Razin, Santiago Oliveros, Michael Richter, Ariel Rubinstein, and two anonymous referees for useful comments. I also thank seminar and conference audiences at the 17th Annual Berkeley/Columbia/Duke/MIT/Northwestern IO Theory Conference, EARIE 2018, ESSM 2017, ESWC 2020, Frankfurt Behavioral IO Workshop, Israeli IO Day, Essex, IDC Herzliya, Hebrew University, LSE, Manchester, RHUL, Technion, Tel Aviv, Toulouse, UCL, Warwick, and Zurich. I thank the Henry Crown Institute of Business Research in Israel and the Solomon Lew Center for Consumer Behavior for financial support.

[1]Membership in the MLM industry is substantial and growing. For example, the global force of independent distributors reached nearly 117 million in 2017 (World Federation of Direct Selling Associations (2018)).

Available at https://econtheory.org. https://doi.org/10.3982/TE4890

a particular company is a legitimate one, or whether it is an exploitative pyramid scam that promotes useless products and services to disguise itself as a legitimate firm, can be a daunting task. One obstacle is that MLM companies typically sell products whose quality is difficult to assess, such as vitamins and nutritional supplements. The common wisdom among practitioners is that a company is legitimate if it encourages the distributors to sell the product, and it is an illegal pyramid scam if it prioritizes recruitment over selling (SEC (2013)). However, it is extremely difficult to determine the company's true "selling point" and, in practice, it is challenging to distinguish between sales to members and sales to the general public.

The objective of this paper is to draw the boundary between the two phenomena on the basis of their underlying reward schemes. Our premise is that the potential distributors are strategic, and that the MLM company chooses a reward scheme while taking these prospective distributors' incentives into account. We refer to reward schemes that compensate distributors for their own sales (of the product and/or distribution licenses) as well as for the sales of the people they recruit as *multilevel schemes*, and to reward schemes that compensate distributors only for their own sales (of the product and/or distribution licenses) as 1-*level schemes*. In principle, it is possible to incentivize distributors to recruit new distributors and sell the product by means of a 1-level reward scheme (e.g., by paying them a recruitment commission for every person they recruit and a sales bonus for every unit they sell). However, in reality, over 90% of the network marketing industry uses multilevel schemes (Direct Selling Association (2014)). Moreover, although there is no obvious reason why 1-level schemes cannot be used for the purpose of sustaining a pyramid scam, various companies that were deemed pyramid scams used multilevel schemes.[2] In this paper, we suggest a mechanism that can explain these stylized facts.

To better understand the difference between MLM and pyramid scams, we develop a model in which a scheme organizer (SO) sells a good to a network of agents that is formed randomly and sequentially, where the agents buy only from people to whom they are directly connected. To reach new pools of customers (agents to whom the SO is not directly connected), the SO uses a reward scheme to incentivize agents to sell the good and recruit (sell distribution licenses to) new distributors. A key feature of the model is that each agent's likelihood of meeting new potential buyers and distributors decreases over time, which makes it unattractive to join the sales force late in the game.

To capture the idea that the main product that is being traded in a pyramid scam is the right to recruit other agents to the pyramid, assume for a moment that the good has no intrinsic value such that the only "products" being traded in the model are distribution licenses. If there exists a reward scheme such that the SO makes a strictly positive profit in its induced game, then we have a *pyramid scam*. Classic no-trade theorems rule out such scams for rational agents in our model, reflecting the fact that such agents cannot be fooled.[3] Hence, to better understand such scams and their underlying reward schemes, we depart from the rational expectations paradigm.

[2]See, for example, *FTC v. Fortune Hi-Tech Marketing Inc.* (2013) for a pyramid scam that used a multilevel reward scheme to recruit over 100,000 distributors in the United States and Canada.

[3]While under the classic rational expectations paradigm agents cannot be scammed, in practice, we observe countless pyramid scams (see, e.g., Keep and Vander Nat (2014), and the references therein).

We use the analogy-based expectation equilibrium (ABEE) framework (Jehiel (2005)) to relax the requirement that each agent has a perfect perception of the other agents' behavior. Under the behavioral model, agents neglect the fact that other agents' strategies are time-contingent. As a result, they underestimate the extent to which recruiting new members becomes more difficult over time. Despite this mistake, each agent's beliefs are statistically correct and can be interpreted as resulting from the use of a simplified model of the other agents' behavior.

We establish that if there are sufficiently many agents, then the SO can sustain a pyramid scam; that is, there are reward schemes that enable the SO to make a strictly positive profit in an ABEE even if the good has no intrinsic value. After establishing the existence of such schemes, we study their structure and show that they (i) charge a license fee and (ii) pay for at least two levels of downline recruits. In other words, there exists no 1-level scheme such that the SO makes a strictly positive profit in its induced game when the only products being traded are distribution licenses.

While the intuition for the existence result is similar to the intuition in previous environments in which ABEE has been used to explain speculation—for example, the centipede game (Jehiel (2005)) or the capped bubble game (Moinas and Pouget (2013))—the intuition for the impossibility result is new. As in previously studied environments, ABEEs have a threshold feature: agents buy licenses up to some time t and then stop. We show that, conditional on buying a license at time t, an agent whose beliefs are statistically correct (as in an ABEE) cannot expect to sell more than one license. Thus, no commission on the agent's own sales (of licenses) would cover the license fee and make it beneficial for him to participate. Note that, in addition to overestimating the number of direct recruits, our agent also overestimates the number of agents his recruits will recruit. When the SO uses a multilevel scheme, these mistakes accumulate, and so the agent may find it worthwhile to purchase a license.

To obtain a better understanding of MLM, we investigate a setting in which the good is intrinsically valued. We solve for the SO's optimal scheme under two behavioral assumptions. First, if the agents are fully rational, then the optimal scheme does not charge license fees, nor does it pay recruitment commissions. Instead, it compensates distributors for the number of units they sell and for the number of units their recruits sell. Second, when agents are analogy-based reasoners, then the properties of the optimal scheme depend on the demand for the good. When the demand is sufficiently high, the optimal scheme looks just like when agents are fully rational. However, when the demand is sufficiently low, the optimal scheme charges a license fee and pays commissions for at least two levels of downline recruits. Thus, the tools that pyramid scams are based on—multiple recruitment commissions and license fees—disappear when the demand for the good is high and emerge again when the demand is low.

We study the implications of banning license fees and recruitment commissions and find that such a regulation may reduce the profit of an SO who faces analogy-based reasoners. However, we show that there is a limit to this negative effect. Even under such a regulation, the SO can still obtain an expected profit that is at least as high as the *fundamental value* of the operation, namely, the expected profit that an unregulated SO could obtain from fully rational agents. Thus, such a regulation could eliminate distributor losses and would never force a business that can survive with rational participants to shut down.

Related literature

Our paper relates to the strand of the behavioral industrial organization literature that studies market settings and contractual features that enable firms to exploit agents who are subject to different biases. Spiegler (2011) offers a textbook treatment of such models and Heidhues and Kőszegi (2018) provide a comprehensive review of this strand of the literature.

In the context of MLM, an independent paper by Stivers, Smith, and Jin (2019) studies important aspects in the relation between MLM schemes and pyramid scams. The authors assume that agents are overoptimistic about the possibility of recruiting new members. As a result of this overoptimism, the agents purchase distributorships regardless of whether they believe that retailing the good is profitable, which results in a pyramid scam. There are two main differences between their model and ours. First, Stivers, Smith, and Jin (2019) take a reduced-form approach to how agents form their beliefs while in the present paper agents' beliefs emerge endogenously in equilibrium. This endogeneity limits the agents' ability to err and allows us to investigate the extent of biases that are needed to support a pyramid scam. Second, Stivers, Smith, and Jin (2019) restrict attention to 1-level reward schemes while the present paper allows for a richer set of schemes. This richness allows us to characterize optimal MLM schemes and explore conditions on the reward scheme that make pyramid scams viable rather than not.

We use the analogy-based expectation equilibrium (Jehiel (2005)) as our main behavioral framework. A closely related concept, the partially cursed equilibrium, was developed by Eyster and Rabin (2005) for Bayesian games. In a partially cursed equilibrium, each agent fails to realize the extent to which the other agents' actions depend on their private information. Similar ideas have been applied in various contexts and, in particular, in the context of behavioral industrial organization. For example, Piccione and Rubinstein (2003) study intertemporal pricing when consumers reason in terms of a coarse representation of the correct equilibrium price distribution.

The pure pyramid scams in our model resemble speculative bubbles (note that bubbles do not include the design and recruiting aspects of a pyramid). Shiller (2015) describes such bubbles as naturally occurring Ponzi processes.[4] Bianchi and Jehiel (2010) and Moinas and Pouget (2013) show that the analogy-based expectation equilibrium logic can sustain a bubble, and Jehiel (2005) shows that it can sustain cooperation in the finite-horizon centipede game, which can be interpreted as a speculative bubble.[5]

[4]Although Ponzi schemes and pyramid schemes are related and these terms are often used synonymously, they are different in several important aspects. In particular, Ponzi scheme participants are not required to recruit new members to make a profit. Moreover, they sometimes believe that an ordinary investment underlies the operation. For example, the participants in Wincapita, a Finnish Ponzi scheme, did not know the true nature of the scheme until it collapsed (Rantala (2019)).

[5]Moinas and Pouget (2013) also provide experimental evidence that the analogy-based expectation equilibrium logic is relevant for speculative bubbles.

Abreu and Brunnermeier (2003) and Moinas and Pouget (2013) study models in which investors become aware of a finite bubble sequentially and face uncertainty about the time at which the bubble started. In both models, a bubble can be sustained in equilibrium under common knowledge of rationality. Unlike in these models, in the present paper classic no-trade arguments (Tirole (1982)) hold and so uncertainty about the time at which the operation started cannot lead to participation in a pure pyramid scam without deviations from the rational expectations paradigm. The reason that no-trade arguments do not hold in the previous models is either that the induced trading game is not a negative-sum game (in Abreu and Brunnermeier's model) or that traders can suffer a (potentially) infinite loss (in Moinas and Pouget's model). In our game, on average, pyramid scam participants incur losses, and their potential loss is bounded.

MLM schemes have received considerable attention outside of the economics literature. A strand of the computer science literature (see, e.g., Emek, Karidi, Tennenholtz, and Zohar (2011), Babaioff, Dobzinski, Oren, and Zohar (2012)) focuses on MLM schemes' robustness to Sybil attacks. The marketing literature has addressed ethical issues in MLM and the resemblance of such schemes to pyramid scams. The common view in that literature is that a company is a pyramid scam if the participants' compensation is based primarily on recruitment rather than sales to end users (see, e.g., Koehn (2001), Vander Nat and Keep (2002)).

The paper proceeds as follows. The model is presented in Section 2. The main analysis is performed in Sections 3 and 4. Section 5 concludes. All formal proofs are relegated to the Appendix.

2. The model

There is a scheme organizer (SO) who produces a good free of cost and with no capacity constraints, and a set of agents $I = \{1, \dots, n\}$. Each agent $i \in I$ has a unit demand and a willingness to pay $q \geq 0$.

In each period $t = 1, 2, \dots, n$, one agent enters the game. We refer to the tth entrant as agent i_t. Upon entering the game, agent i_t meets one player $j \in \{\text{SO}, i_1, \dots, i_{t-1}\}$ chosen uniformly at random by nature.[6] For example, agent i_2 meets either the SO or agent i_1, each with probability 0.5. In period 1, the SO can offer agent i_1 the opportunity to purchase a distribution license. Conditional on receiving an offer, i_1 can accept or reject it. Let $D_1 = \{\text{SO}, i_1\}$ if i_1 accepts an offer and $D_1 = \{\text{SO}\}$ otherwise. In each period $t > 1$, if agent i_t meets a player $j \in D_{t-1}$, then j can offer him the opportunity to purchase a distribution license. Otherwise, if i_t meets an agent $j \notin D_{t-1}$, then he will not be offered the opportunity to purchase a distribution license. If he receives an offer, i_t can accept or reject it. Let $D_t = D_{t-1} \cup \{i_t\}$ if i_t accepts an offer and $D_t = D_{t-1}$ otherwise.

Let G denote the directed tree, rooted at the SO, that is induced by the above process, where each node represents an agent and each meeting is represented by an edge that points away from the root. We use $d(i, j)$ to denote the length of the directed path from $i \in I \cup \{\text{SO}\}$ to $j \in I$ if such a path exists.

[6] We assume that nature makes her choice at the beginning of period t.

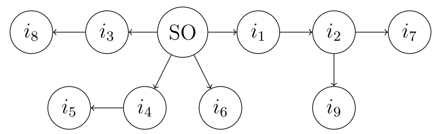

FIGURE 1. A snapshot of G at the end of period 9.

To illustrate the game tree, suppose that, at the end of period 9, G is as presented in Figure 1 and $D_9 = \{\text{SO}, i_1, i_2, i_3\}$. Recall that agent i_{10} is equally likely to meet each $j \in \{\text{SO}, i_1, \dots, i_9\}$. If i_{10} meets $j \in D_9$, then j decides whether or not to make i_{10} an offer. If i_{10} meets an agent $j \notin D_9$, then agent i_{10} does not receive an offer.

Agents who meet a distributor or the SO can purchase the good for personal consumption. The choice of whether to purchase the good or not and the choice of whether to sell the good or not are modeled in a nonstrategic manner. We assume that the SO fixes a price η at the beginning of the game and that when an agent meets a distributor or the SO, he buys a unit of the good for personal consumption if and only if $\eta \leq q$ (regardless of whether he receives an offer to purchase a license or not).

Agents in the model make two types of strategic choices: whether to accept an offer (to purchase a license) upon entering the game, and conditional on purchasing a license, whether to make an offer when they meet agents later in the game. In each period t, each player $i \in \{\text{SO}, i_1, \dots, i_t\}$ knows the period t, his immediate predecessor in G, the realization of the subtree of G rooted at i up to period t, and for every agent j in that subtree, whether j purchased a license. Denote by H_i the set of information sets in which player $i \in I \cup \{\text{SO}\}$ moves. Player i's strategy is a mapping $\sigma_i : H_i \to \{0, 1\}$, where in each $h \in H_i$, i chooses whether or not to make an offer, or else chooses whether or not to accept one. We use $\sigma = (\sigma_i)_{i \in I \cup \{\text{SO}\}}$ to denote a profile of strategies.

The transfers between the SO and the distributors are determined according to a reward scheme that consists of

- A license fee ϕ.
- Recruitment commissions: $a_1, a_2, a_3, a_4, \dots \geq 0$.
- Sales bonuses: $b_1, b_2, b_3, b_4, \dots \geq 0$.

The SO receives ϕ from every agent who purchases a distribution license (i.e., accepts an offer) and η from every agent who purchases the good for personal consumption, and pays recruitment commissions and sales bonuses to distributors based on their success and their downline recruits' success in recruiting new distributors and selling the good. Formally, when an agent j accepts an offer for a distribution license, he pays ϕ to the SO and the SO pays a recruitment commission of $a_{d(l,j)}$ to every distributor l on the directed path from the SO to j. When an agent j purchases the good, he pays η to the SO, and the SO pays a sales bonus of $b_{d(l,j)}$ to every distributor l on the directed path from the SO to j. Let $\mathbf{a} := (a_\tau)_{\tau \geq 1}$, $\mathbf{b} := (b_\tau)_{\tau \geq 1}$, and for every integer z, let $\mathbf{a}_{-z} := (a_\tau)_{\tau > z}$, and $\mathbf{b}_{-z} := (b_\tau)_{\tau > z}$.

We say that a reward scheme is a *z-level scheme* if $\mathbf{a}_{-z} = \mathbf{0}$ and $\mathbf{b}_{-z} = \mathbf{0}$, and $a_z \neq 0$ or $b_z \neq 0$. For example, in a 1-level scheme $\mathbf{a}_{-1} = \mathbf{b}_{-1} = \mathbf{0}$, and so the agents are compensated only for their direct sales and recruitments. In a 2-level scheme, agents are also compensated for their recruits' sales and recruitments.

We shall focus on schemes where $a_\tau \leq \phi$ and $b_\tau \leq \eta$ for every $\tau \geq 1$, and refer to such schemes as *incentive-compatible* (IC) schemes. This constraint implies that for a distributor, the cost of creating a fictitious new tree of sales and recruits is greater than the direct benefit of doing so (i.e., the transfers from the SO to the root).

The payoff of an agent who purchases a license consists of the transfers specified by the reward scheme and two additional costs. First, when an agent accepts an offer, he incurs a cost of $c > 0$, which reflects the cost of learning how to market the good. Second, the person who recruited him incurs a cost of $\hat{c} > 0$, which reflects the cost of training the new distributor.[7] We assume that when the SO recruits an agent he also incurs a cost of $\hat{c}$.

We denote by $\Gamma(R)$ the game that is induced by the reward scheme R. Agents are risk neutral and maximize their expected payoff in $\Gamma(R)$ given the reward scheme and their beliefs about other players' behavior. The highest expected payoff the SO obtains across all equilibria of $\Gamma(R)$ is denoted by $\pi(R)$. A scheme R is said to be profit-maximizing if there exists no scheme R' such that $\pi(R') > \pi(R)$.

2.1 *Discussion: Modeling assumptions*

Meeting process We borrow the meeting process from the applied statistics literature, where it is referred to as the *uniform random recursive tree model* (for a textbook treatment, see Drmota (2009)).[8] This process rests on the assumptions that there is a deterministic date at which the game ends, that only one meeting occurs in each period, and that each agent is directly connected to only one agent upstream (i.e., agents have market power in their social neighborhood). Our main results and insights do not depend on the specific network formation process or its underlying assumptions, as we shall explain in Section 5. We use this process since it allows us to convey the main messages while keeping the exposition relatively simple.

Incentive compatibility The SO faces the risk that the distributors will create fictitious recruits to become eligible for additional commissions and bonuses.[9] Motivated by this risk, we shall focus on IC schemes. This constraint mitigates the above risk since it implies that, for a distributor, the cost of creating a fictitious new tree of sales and recruits is greater than the direct benefit of doing so (i.e., the transfers from the SO to the root). The incentive-compatibility constraint rests on the assumption that the SO can verify

[7] It is possible to incorporate moral hazard into the model by adding a choice of whether to train the new recruit without changing the key results of the paper.

[8] Gastwirth (1977) and Bhattacharya and Gastwirth (1984) used this model to examine two real-world pyramid scams and to demonstrate that only a small fraction of the participants can cover the license fees. Neither of these papers, however, takes strategic considerations into account.

[9] In the computer science literature, manipulations in this spirit are often referred to as local false-name manipulations or local splits (see, e.g., Emek et al. (2011), Babaioff et al. (2012)).

the identity of any distributor who wishes to receive commissions and, therefore, even if a distributor were to create a fictitious recruit, he would not be able to collect the commissions that the fictitious recruit is eligible to receive.[10]

2.2 *Benchmark: The social optimum*

From a social perspective, the cost of turning agent i_t into a distributor is $c+\hat{c}$: c is incurred by i_t when learning how to market the good, and $\hat{c}$ is incurred by his recruiter when training i_t. Recall that agents consume the good if and only if they meet a distributor upon entering the game. In expectation, agent i_t meets

$$v_t := E\big[\big|\big\{j \in I : d(i_t, j) = 1\big\}\big|\big] = \frac{1}{t+1} + \cdots + \frac{1}{n} \tag{1}$$

new entrants in periods $t+1, \dots, n$ (for completeness, let $v_n := 0$). Thus, if agent i_t becomes a distributor he provides the good to v_t agents who would not consume it otherwise. Hence, the direct social benefit from turning agent i_t into a distributor is qv_t. If $qv_t > c+\hat{c}$, then it is socially desirable that agent i_t will purchase a license even if we do not take into account downline recruitment. Since v_t is monotonically decreasing in t, if $qv_t < c+\hat{c}$, then from a social perspective it is best if agent i_t does not recruit anyone. We can conclude that if $qv_t < c+\hat{c}$ then it is socially undesirable that i_t will obtain a license even if downline recruitment is taken into account. Hence, the social optimum is characterized by a cutoff period $k^w = \sup\{t | qv_t \geq c+\hat{c}\}$ such that agents obtain distribution licenses if and only if they enter the game before k^w.[11]

We say that a reward scheme R is socially optimal if there exists an equilibrium of $\Gamma(R)$ that implements a socially optimal allocation of distribution licenses, that is, an equilibrium in which, for every $t \in \{1, \dots, n\}$, agent i_t purchases a license if and only if

$$qv_t \geq c+\hat{c}. \tag{2}$$

3. Fully rational agents

Our objective in this section is to understand the properties of reward schemes that maximize the SO's expected payoff when he cannot scam the agents. To this end, we assume that agents are fully rational, and hence not vulnerable to deceptive practices. In this setting, the reward scheme incentivizes agents to sell the good and recruit others to the sales force, which allows the SO to reach potential customers to whom he is not directly connected who would not buy the good otherwise.

To capture that the agents are fully rational, we use perfect Bayesian equilibrium (PBE) to solve the model. Throughout the analysis, we assume that an agent who is indifferent whether to accept an offer or not accepts it, and that a distributor who is

[10]We shall discuss the verifiability assumption in detail in the concluding section.

[11]In the optimal stopping literature, such rules are known as "one-step look-ahead rules" or "myopic rules."

indifferent whether to make an offer or not makes it.[12] Finally, since the agents' behavior is independent of η for all prices $\eta \leq q$, we shall fix $\eta = q$ without loss of generality.

We start the analysis with a simple benchmark result that establishes that indeed the SO cannot scam fully rational agents.

3.1 *Pure pyramid scams ($q = 0$)*

To capture the idea that the main "good" being traded in a pyramid scam is the right to recruit others to the pyramid, we set $q = 0$, in which case it is commonly known that the only goods that are being traded in the model are distribution licenses. Intuitively, such a market should not exist as trade in distribution licenses does not add any value. Indeed, at the social optimum, no agent obtains a license in this case. We say that the SO sustains a *pure pyramid scam* if $q = 0$ and there exists a reward scheme such that the SO makes a strictly positive expected payoff in an equilibrium of its induced game. Proposition 1 establishes that when agents are fully rational it is impossible for him to do so.

Proposition 1. *Let $q = 0$. There exists no scheme R such that $\pi^{\text{PBE}}(R) > 0$.*

When $q = 0$, reward schemes induce negative-sum transfers between the agents and the SO. Proposition 1 then follows from classic no-trade arguments (Tirole (1982)).[13]

3.2 *Multilevel marketing of genuine goods ($q > 0$)*

Having established that the SO cannot make a profit when he produces worthless products, we turn to the case in which the agents are willing to pay for the product and study their equilibrium behavior and the schemes that maximize the SO's expected payoff. Let us start with the agents' perspective. Since the likelihood of meeting new entrants goes down over time, agents who enter the game in its early stages meet more people, and so have more opportunities to sell the good and recruit new members. As the cost of purchasing a license is fixed, early entrants find it more beneficial to purchase a license than later entrants. Moreover, since the cost of training a new recruit is fixed, distributors find it more beneficial to recruit a new entrant in the early stages of the game rather than in its later stages (we say that a distributor recruits an agent if he makes an offer to the latter and the latter accepts). The next lemma formalizes this argument and shows that the equilibrium of the game has a threshold structure.

Lemma 1. *Consider a scheme R. In every PBE of $\Gamma(R)$:*

(i) If agent i_t receives an offer, he accepts it if and only if $v_t b_1 \geq c + \phi$.

[12] Our results are not sensitive to these assumptions since agents break their indifference in favor of accepting and making offers in the PBE that maximizes the SO's expected payoff.

[13] In fact, while the present proof relies on the sequential rationality property of PBE, the impossibility result is more general and can be extended to Bayesian Nash equilibria.

(ii) *A distributor who meets agent i_t recruits him if and only if $v_t b_1 \geq c + \phi$ and $v_t b_2 + a_1 \geq \hat{c}$.*

Lemma 1 establishes that the agents' behavior is identical in every PBE. To obtain intuition for the first condition in Lemma 1, note that $v_t b_1$ is the bonuses agent i_t expects to obtain from the sales he makes *directly* and $c + \phi$ is the total cost of becoming a distributor. Thus, when $v_t b_1 \geq c + \phi$, agent i_t finds it optimal to purchase a license even without taking the possibility of recruiting new members into account. To see why $v_t b_1 \geq c + \phi$ is also a necessary condition for i_t to purchase a license, note that, if i_t is the last agent who is supposed to purchase a license in a PBE, he cannot expect to recruit anyone. He thus finds it optimal to purchase a license only if $v_t b_1 \geq c + \phi$. Since v_t is decreasing in t, this condition must hold for every agent who purchases a license before t, namely, every agent who purchases a license in equilibrium.

To understand the second condition in Lemma 1, note that a distributor who recruits agent i_t obtains, in expectation, $v_t b_2$ for i_t's direct sales and a_1 for recruiting i_t. Thus, if $v_t b_2 + a_1 \geq \hat{c}$, then it is optimal for distributors to recruit i_t regardless of whether they believe that later entrants will purchase licenses or not. To see that this condition is also necessary, note that it must hold for the last period t in which a distributor is supposed to recruit an entrant as that distributor cannot expect the entrant to recruit anyone. Since v_t is decreasing in t, $v_t b_2 + a_1 \geq \hat{c}$ for every period t in which distributors are supposed to recruit in equilibrium.

The conditions in Lemma 1 imply that the agents' behavior in equilibrium is characterized by two thresholds, k_1 and $k_2 \leq k_1$, such that every agent who enters the game up to period k_1 accepts every offer he receives, and distributors recruit every agent they meet up to period k_2. Note that if $k_2 < k_1$, then agents may receive offers in periods $k_2 + 1, \ldots, k_1$ only if they meet the SO. The SO makes an offer in period $t \in \{k_2 + 1, \ldots, k_1\}$ if the cost of training the new recruit $\hat{c}$ is less than the expected net revenue from the new recruit's direct sales $v_t(q - b_1)$ as the new recruit is not expected to recruit new recruits himself.

Having characterized the structure of PBE, we now turn to study the scheme that maximizes the SO's expected payoff. By Lemma 1, the agents' PBE behavior depends only on ϕ, b_1, a_1, and b_2. Thus, schemes that pay higher-order commissions can only increase the SO's cost and reduce his profit. Therefore, in the remainder of this section, we shall focus on 2-level schemes with $a_2 = 0$.

Theorem 1.

- *If $q > \frac{c+\hat{c}}{v_2} + \frac{c}{2v_2^2}$, then the profit-maximizing scheme sets $(\phi, a_1, b_1, b_2) = (0, 0, \frac{c}{v_{k_1}}, \frac{\hat{c}}{v_{k_2}})$, where the cutoff periods satisfy $k_1 \geq k_2 \geq 2$. Moreover, at least two agents purchase a license in every PBE of its induced game.*
- *If $q \in (\frac{c+\hat{c}}{v_1}, \frac{c+\hat{c}}{v_2} + \frac{c}{2v_2^2})$, then every scheme that pays $b_1 = \frac{c+\phi}{v_1}$ is profit-maximizing. Moreover, in every PBE of a game induced by such a scheme exactly one agent purchases a license.*

- *If $q < \frac{c+\hat{c}}{v_1}$, then every scheme is profit-maximizing. Moreover, in every PBE, none of the agents purchase a license.*

The condition $q > (c+\hat{c})/v_2 + c/(2v_2^2)$ guarantees that the agents' willingness to pay q is sufficiently large relative to n, c, and $\hat{c}$ such that, at the optimum, the SO incentivizes (at least) two agents to purchase a license. When this condition is violated, incentivizing multiple agents to purchase a license is suboptimal, which renders MLM irrelevant.

To see why charging a license fee is detrimental to the SO when MLM is relevant, consider the last agent who is supposed to purchase a license in a PBE. At the optimum, this marginal agent is indifferent whether to purchase a license or not as, otherwise, the SO could charge a higher license fee without affecting the agents' PBE behavior. The fee ϕ must be compensated by direct sales bonuses for the marginal agent. Since inframarginal agents earn more bonuses, a reduction in ϕ and b_1 that keeps the marginal agent indifferent reduces the SO's cost. In other words, scaling down ϕ and b_1 reduces the rents agents obtain by purchasing a license early in the game.

Note that IC schemes that do not charge a license fee cannot pay directly for recruitment; that is, $\phi = 0$ implies that $a_1 = 0$. Thus, by refraining from charging a fee the SO loses some flexibility in choosing how to incentivize distributors to recruit new members: he must use b_2 rather than a_1 or a combination of the two. As we show in the proof, this effect is of second order such that the SO is better off setting $a_1 = \phi = 0$.

Lemma 1 implies that if k_1 is the last period in which agents accept offers and k_2 is the last period in which distributors make offers, then $b_1 v_{k_1} \geq c + \phi$ and $a_1 + b_2 v_{k_2} \geq \hat{c}$. Since a profit-maximizing scheme must leave the marginal agent and the marginal distributor indifferent, $(a_1, \phi) = (0, 0)$ implies that $(b_1, b_2) = (c/v_{k_1}, \hat{c}/v_{k_2})$.

We now revisit the restrictions on the contract space and explore how they enable the distributors to obtain rents for purchasing a license early in the game. First, the restriction to time-invariant schemes prevents the SO from discriminating between the distributors based on the time at which they purchase a license. If the SO could charge a fee that decreases over time, then he would be able to charge each distributor a fee that leaves the latter with no surplus without affecting the other agents' behavior. This, in turn, would allow the SO to maximize the social surplus and capture all of it.[14] Second, the incentive-compatibility restriction prevents the SO from directly compensating the agents for their costs. For example, a non-IC 1-level scheme that charges $\phi = -c$ and pays $a_1 = \hat{c}$ and $b_1 = 0$ would compensate the agents' directly for every cost incurred, leave them with no rents, and implement the social optimum.

In general, profit-maximizing schemes are not socially optimal: they pay lower commissions and incentivize fewer agents to become distributors than is socially optimal. Under a socially optimal scheme $k_1^w = k_2^w =: k^w$, whereas under profit maximization $k_2 \leq k_1 \leq k^w$ (with strict inequalities when n and q are not too small relative to c and $\hat{c}$). As a result, profit-maximizing schemes result in fewer agents buying the good than is socially optimal. To gain intuition, note that, from the SO's perspective, raising the

[14]For example, the SO could use a 1-level scheme that charges each distributor i_t a fee of $\hat{c} + \max\{(v_t - v_{k^w})(c+\hat{c})/v_{k^w}, 0\}$ and pays a fixed bonus of $b_1 = (c+\hat{c})/v_{k^w}$ and a fixed commission of $a_1 = \hat{c}$. In a PBE of the induced game, the first k^w entrants purchase a license and receive an expected payoff of 0.

number of distributors from $t-1$ to t entails paying higher bonuses that make it worthwhile for i_t to buy a license. Since all agents are paid according to the same scheme, by increasing the bonuses the SO essentially forgoes some of the profit from the first $t-1$ distributors' sales. The social point of view ignores this effect and, therefore, socially optimal schemes induce more distributors relative to profit-maximizing schemes.

4. Boundedly rational agents

Our main objective in this section is to understand the forces and compensation plans that enable pyramid scams to operate. As Proposition 1 suggests, we cannot do so by means of the classic rational expectations model and we shall therefore depart from this model. We shall weaken the Nash equilibrium assumption that agents have perfect perception of the other agents' behavior in every possible contingency, an assumption that might be too extreme when individuals face complicated new situations.

Jehiel (2005) suggests an elegant framework that incorporates partial sophistication into extensive-form games. In this framework, different contingencies are bundled into analogy classes and the agents are only required to hold correct beliefs about the other agents' *average behavior* in every analogy class.

Our agents have this type of correct, yet coarse, perception of the other agents' behavior. They understand the frequencies at which the other agents accept and make offers. However, they do not understand the time-contingent nature of the other agents' behavior. In simple words, agents do not base their expectations that offers will be accepted on the time at which they are made. Instead, they pool all offers made at any point in time and consider the average rate of offer acceptances. Thus, agents view other agents' behavior as if it were time-invariant and underestimate the extent to which it becomes more difficult to recruit new members over time.

In equilibrium, the agents' beliefs about the other agents' behavior are statistically correct, and can be interpreted as a result of learning from partial feedback about the behavior in similar games that were played in the past (e.g., similar schemes organized by the SO). One motivation for the agents' coarse reasoning is that obtaining feedback about the aggregate behavior in these past schemes' induced games might be easier than gathering information about the time and context in which each offer was made.

Under the solution concept, agents' beliefs about nonstrategic aspects and, in particular, the network formation process are correct. Thus, the solution concept does not distort agents' beliefs about v_t.

Formally, for each $i \in I$ we denote by H_i^1 the set of information sets in which i chooses whether or not to purchase a license, and by H_i^2 the set of information sets in which i chooses whether or not to make an offer. Let $M_1 := \bigcup_{i\in I} H_i^1$ and $M_2 := \bigcup_{i\in I} H_i^2$. We refer to M_1 and M_2 as the agents' analogy classes and denote by $r_\sigma(h)$ the objective probability of reaching $h \in M_1 \cup M_2$ conditional on the profile σ being played. For each $i \in I$, $\alpha^i = (\alpha^i(M_1), \alpha^i(M_2))$ are agent i's analogy-based expectations about the other agents' behavior. A strategy σ_i is a best response to α^i if it is optimal given a belief that each agent $j \neq i$ accepts every offer he receives with probability $\alpha^i(M_1)$ and that, if j has the opportunity to make an offer, then he makes it with probability $\alpha^i(M_2)$. Let $\alpha := (\alpha^i)_{i\in I}$.

Definition 1. Agent i's analogy-based expectations α^i are said to be consistent with the profile of strategies σ if, for every $k \in \{1, 2\}$, it holds that

$$\alpha^i(M_k) = \frac{\sum_{h \in M_k} r_\sigma(h)\sigma(h)}{\sum_{h \in M_k} r_\sigma(h)}$$

whenever $r_\sigma(h) > 0$ for some $h \in M_k$.

Definition 2. The pair (σ, α) forms an analogy-based expectation equilibrium (ABEE) if, for each $i \in I$, α^i is consistent with σ and σ_i is a best response to α^i.

Consistency implies that, in an ABEE, $\alpha^i(M_1) = \alpha^j(M_1)$ and $\alpha^i(M_2) = \alpha^j(M_2)$ for every pair of agents $i, j \in I$. Therefore, we shall omit the superscript in the remainder of the paper. Moreover, for the sake of brevity, we shall omit the dependence on M_1 and M_2 and use α_1 and α_2 instead of $\alpha(M_1)$ and $\alpha(M_2)$, respectively.

Discussion: Analogy classes, consistency, and the SO's strategy

Analogy classes Each agent i's analogy classes, M_1 and M_2, consist of all of the information sets in which agents move, including information sets in which i himself moves.[15] This is consistent with interpreting i's behavior as best responding to coarse feedback about the behavior in similar games that were played in the past by a different set of agents (i.e., i himself did not play in these games). Note that since i was not a player in these past games, his own actions do not affect his analogy-based expectations.

We could exclude the information sets in which agent i moves from his own analogy classes. These alternative analogy classes are consistent with the interpretation of i's behavior as best responding to coarse feedback about the behavior in similar games in which i himself played in the past. Our results hold under both types of partitions.[16]

Consistency Definition 1 does not place any restrictions on the agents' beliefs about analogy classes that are not reached with strictly positive probability. We can refrain from placing such restrictions as the only equilibria in which M_1 and M_2 are not reached with strictly positive probability are equilibria in which the SO never makes any offers, which are of secondary interest and do not change our results.

Consistency implies that the agents' expectations α_1 match the proportion of accepted offers. An important feature of consistency is that information sets are weighted according to the likelihood of their being reached. To see this, let $n = 3$ and consider a profile σ in which the SO makes an offer in period 1 and, in each period $t \in \{2, 3\}$, every

[15] We refrain from defining an analogy class for the agents' decisions on whether to purchase the good or not (which are nonstrategic). Were we to define such an analogy class, it would not affect the analysis since the probability with which agents purchase the good is constant over time (and so, in an ABEE, the agents' expectations about the possibility of selling the good would be correct).

[16] In fact, the alternative analogy classes were employed in an earlier version of this paper (Antler (2018)).

$i \in D_{t-1}$ makes an offer if he meets an agent. Moreover, suppose that agent i_1 accepts the SO's offer and all other agents reject every offer they receive. Note that agents i_1 and i_2 always receive an offer under σ. Agent i_3 receives an offer with probability $\frac{2}{3}$ since, with probability $\frac{1}{3}$, he meets agent i_2 who does not have a license. Only the first of the $\frac{8}{3}$ offers is accepted. Hence, $\alpha_1 = \frac{1}{1+1+\frac{2}{3}} = \frac{3}{8}$ is consistent with σ.

The SO's strategy The solution concept does not require that the SO's strategy be optimal. Thus, effectively, the SO is allowed to commit to a strategy. He can potentially benefit from such a commitment as his behavior affects α. The SO's commitment power allows us to simplify the exposition, but does not affect the results.

4.1 *Pure pyramid scams ($q = 0$)*

As in Section 3, we start the analysis with the case of $q = 0$ in which it is commonly known that the players are trading only distribution licenses. Since there is no demand for the good, we shall set $\mathbf{b} = \mathbf{0}$ and $\eta = 0$ throughout this subsection without loss of generality. We focus on the case of $q = 0$ to simplify the exposition. However, as we shall clarify later, all of the results and insights of this section apply when $q > 0$ is small relative to c and $\hat{c}$.

Before proceeding to the analysis, it is useful to mention that when $q = 0$, there always exists a "nonpyramidal" ABEE in which no agent ever purchases a license. To see this, note that given a belief that distributors will be unable to recruit new members ($\alpha_1 = 0$), not buying a license is a best response.

Our main objective in this subsection is to understand the necessary conditions for a pyramid to operate. We first show that the SO cannot sustain a pure pyramid scam by means of 1-level scheme (Theorem 2). Thus, behind every pure pyramid scam in our model is a scheme that pays for at least two levels of downline recruits. Theorem 3 shows that if the number of potential participants is sufficiently large, it is possible to sustain a pure pyramid scam by means of a 2-level scheme, and establishes necessary conditions for such a scheme to be profitable. Theorem 4 shows that the SO may have to use a 3-level scheme to sustain a pure pyramid scam. The analysis in this section establishes that increasing the number of levels increases the agents' overestimation of the value of a distributorship, which in turn can increase the profits of scams and be essential to making them viable.

We now show that the SO cannot sustain a pure pyramid scam by means of a scheme that compensates distributors only for the number of agents they directly recruit.

THEOREM 2. *Let $q = 0$. There exists no 1-level scheme R such that $\pi^{\mathrm{ABEE}}(R) > 0$.*

To gain intuition for this result, let us suppose for a moment that agents purchase licenses in an ABEE. Note that the number of licenses sold is equal to the number of distributors. The SO sells licenses to some of the distributors and, therefore, on average, each distributor sells less than one license. In an ABEE, each agent's beliefs about other agents' tendency to accept offers as well as his beliefs about the number of opportunities to recruit new members are statistically correct, and so the average distributor analogy-based expects to sell less than one license. The incentive-compatibility assumption implies that $a_1 \leq \phi$, and so the average distributor cannot recoup the license fee and, therefore, finds it suboptimal to purchase a license.

Next, we establish necessary conditions for a 2-level scheme to be profitable.

Theorem 3. *Let $q = 0$. There exists an integer $n_2^\star$ such that a 2-level scheme R for which $\pi^{\text{ABEE}}(R) > 0$ exists if and only if $n \geq n_2^\star$. Moreover, every 2-level scheme R that induces a payoff $\pi^{\text{ABEE}}(R) > 0$ for the SO pays $a_1 > 0$, $a_2 > 0$, and $a_1 + a_2 > \phi$.*

The intuition for the existence result resembles the intuition for the ABEE analysis of the finite-horizon centipede game (Jehiel (2005)) and the capped bubble game (Moinas and Pouget (2013)). In an ABEE in which agents accept offers, they accept offers up to some period t and reject offers afterward. However, analogy-based reasoners view the other agents' behavior as if it were time-invariant: each agent falsely believes that each of the other agents accepts offers with probability α_1, even after period t. Thus, agent i_t falsely expects that he will recruit new entrants, that his recruits will recruit new entrants, and so on. This overoptimistic belief is what makes agent i_t pay the license fee and join the pyramid.

What is the difference between multilevel schemes and 1-level schemes? Multilevel and, in particular, 2-level schemes induce contracts that require prospective participants to assess not only the number of people they will recruit in the future but also the number of people their recruits will recruit in the future. The agents' imperfect perception of the other agents' behavior leads the last agents who join the scheme to overestimate both of these variables. As we showed in the proof of Theorem 2, the last agent to purchase a license does not overestimate his own ability to recruit by much and, therefore, the SO cannot overcome the incentive-compatibility constraint and sustain a pyramid scam by means of a 1-level scheme. In a similar manner, this agent does not overestimate his recruits' ability to sell licenses by much. However, the accumulation of mistakes allows the SO to overcome the incentive-compatibility constraint and sustain a pyramid scam.

To see why $a_1 + a_2 > \phi$ is a necessary condition for a 2-level scheme to be profitable, recall from Theorem 2 that the average agent cannot expect to sell more than one license. Since agents who enter the game later have fewer opportunities to sell licenses, the average agent cannot expect that, on average, his successors in the game tree will sell more than one license each. Thus, he cannot expect to have more than one distributor in the second level of his downline. Therefore, he cannot expect a payoff greater than $a_1 + a_2 - c - \phi$, and so finds it suboptimal to purchase a license if $a_1 + a_2 \leq \phi$.

The next example further illustrates the difference between 1- and 2-level schemes by considering a profile of strategies and showing why it cannot be part of an ABEE of a 1-level scheme's induced game but may be part of an ABEE of a 2-level scheme's induced game.

Example 1. Let $n = 100$ and consider a profile σ in which agents accept every offer up to period k_1, reject every offer after period k_1, distributors make an offer to every agent they meet, and the SO makes offers to every agent he meets up to period k_1 and does not make offers afterward. The proof of Theorem 3 shows that σ is consistent with the analogy-based expectations $\alpha_1 = k_1/(k_1 + k_1 v_{k_1})$ and $\alpha_2 = 1$. We now consider different schemes and ask whether σ can be part of an ABEE of their induced games.

First, to illustrate the impossibility result, consider a 1-level scheme R and the marginal agent i_{k_1}. Since $a_1 \leq \phi$, he can recoup the license fee only if he recruits more than one agent. In expectation, he makes v_{k_1} offers and analogy-based expects each of them to be accepted with probability α_1. Since $\alpha_1 v_{k_1} < 1$, agent i_{k_1} finds it suboptimal to purchase a license, in contradiction to σ being part of an ABEE of $\Gamma(R)$.

Second, to illustrate the possibility result, consider a 2-level scheme R. Note that agent i_{k_1} may recoup the license fee if the sum of the number of agents he recruits and the number of people they recruit exceeds 1. He analogy-based expects this sum to be

$$\alpha_1 v_{k_1} + \alpha_1^2 \alpha_2 \sum_{j=k_1+1}^{n} \frac{1}{j} v_j = \frac{k_1}{k_1 + k_1 v_{k_1}} v_{k_1} + \left(\frac{k_1}{k_1 + k_1 v_{k_1}}\right)^2 \sum_{j=k_1+1}^{n} \frac{1}{j} v_j, \tag{3}$$

where $\alpha_1 v_{k_1}$ is the number of people he analogy-based expects to recruit and $\alpha_1^2 \alpha_2 \sum_{j=k_1+1}^{n} v_j/j$ is the number of people he analogy-based expects that these recruits will recruit. As illustrated by the dashed curve in Figure 2, expression (3) is greater than 1 if $k_1 \leq 5$. Thus, potentially, σ can be part of an ABEE of $\Gamma(R)$. For example, if the costs c and $\hat{c}$ are negligible, σ (with $k_1 = 5$) is part of an ABEE of a 2-level scheme that pays $a_1 = a_2 = 0.9853$ and charges $\phi = 1$. That ABEE induces an expected payoff of 2.4546 for the SO. To verify that σ is part of an ABEE, note that (3) is equal to $1/0.9853$ when $k_1 = 5$ such that the marginal agent i_5 is indifferent whether to purchase a license. ◊

In light of Theorem 3, it is natural to ask whether there are cases in which the SO must use more than two levels of compensation. The next result provides an affirmative answer by showing that the required network size for exploiting the agents is smaller for 3-level schemes than it is for 2-level schemes.

Theorem 4. *There exists an integer $n_3^\star < n_2^\star$ such that, for every $n \geq n_3^\star$, there exists a 3-level scheme R such that $\pi^{\text{ABEE}}(R) > 0$.*

Theorem 4 shows that the SO can strictly benefit from paying higher-order commissions such as a_3 when he faces analogy-based reasoners. This commission raises the marginal agent's willingness to pay for a license as he falsely believes that his recruits may recruit new distributors who, in turn, may recruit new distributors themselves. This false belief is illustrated in the difference between the solid and dashed curves in Figure 2. Due to this false belief, increasing the number of levels of the scheme increases the marginal agent's overestimation of the value of a distributorship, which makes the scam viable when $n_3^\star \leq n < n_2^\star$.

Theorem 4 illustrates a more general phenomenon: the minimal number of agents that is required to sustain a pyramid scam weakly decreases with the number of levels of

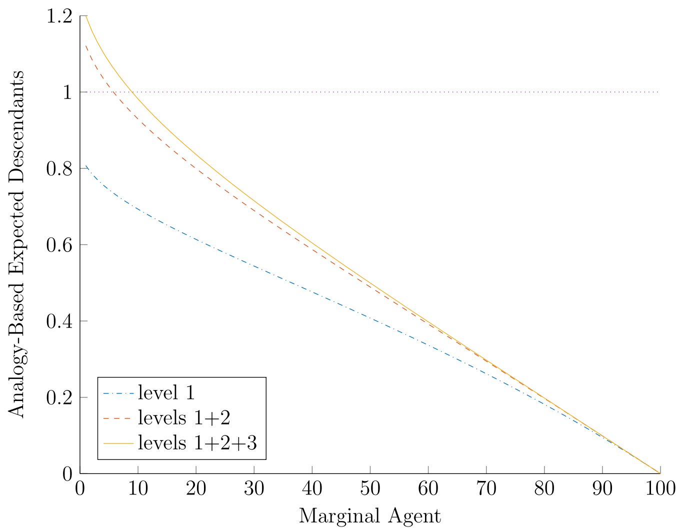

Figure 2. The analogy-based number of members in the first, second, and third levels of the marginal agent's downline in Example 1.

the scheme that the SO is allowed to use. The same techniques used in the proof of the theorem can be used to show that $n_4^\star$, the minimal number of agents required to sustain a pyramid scam using a 4-level scheme, is strictly smaller than $n_3^\star$. In general, $n_k^\star$ is only weakly decreasing in k since it is bounded from below.

By contrast, in our benchmark model in which the SO faces fully rational agents he cannot benefit from paying higher-order commissions. These commissions are not beneficial since the fully rational agents correctly predict the other agents' behavior. In particular, the marginal fully rational agent understands that he will not recruit anyone and, therefore, cares only about the direct sales bonus b_1 when considering taking an offer (see Condition (i) in Lemma 1). Similarly, the marginal fully rational distributor correctly expects that a new recruit will not recruit anyone and, therefore, cares only about a_1 and b_2 when considering making an offer (see Condition (ii) in Lemma 1).

In conclusion, the results in this section show that if the network size is sufficiently large, then the SO can sustain a pyramid scam. Sustaining a scam requires a scheme that charges a license fee and pays for recruitment, as the good is not intrinsically valued. Moreover, it requires paying for at least two levels of downline recruits and, in some instances, for strictly more than two levels. These results are different from the case in which agents are fully rational, where the SO never finds it beneficial to charge license fees, pay recruitment commissions, or use schemes that pay for more than two levels of downline sales.

Our analysis focused on time-invariant schemes. Such schemes are relatively simple to explain to potential participants and may conceal the nonstationary nature of the environment from these potential participants. We now relax the stationarity assumption and illustrate that using a scheme with license fees and recruitment commissions that weakly rise over time is another way for the SO to target the agents' mistaken beliefs.

Specifically, we show that the SO can sustain a pyramid scam by means of a nonstationary 1-level scheme. To adapt the reward scheme, assume that an agent who buys a license in period t pays ϕ^t and that a distributor who recruits in period t is paid a_1^t. We impose that $a_1^t \leq \phi^t$ to adapt the incentive-compatibility constraint. Let $\phi^t = \hat{c} + \epsilon$ and $a_1^t = \hat{c}$ for $t < n$ and $\phi^n = a_1^n = m$. Consider a profile σ in which distributors always make offers and agents accept every offer up to period n and reject offers made in period n. This profile is consistent with $\alpha_1 = (n-1)/n$. Thus, agent i_t (where $t < n$) analogy-based expects a payoff of $-c - \hat{c} - \epsilon + (m(n-1))/((n+1)n)$ from purchasing a license. If m is large enough, σ is part of an ABEE in which the SO's profit is $(n-1)\epsilon > 0$.

To gain intuition, note that when the fee increases over time, an agent can earn commissions greater than the fee that he pays without violating the incentive-compatibility constraint (in the example, $a_1^n > \phi^t$ for $t < n$). Thus, an agent may find it beneficial to purchase a license even if he does not expect to recruit more than one person, which is a necessary condition for participation when a 1-level scheme is stationary.

4.2 *Multilevel marketing of genuine goods ($q > 0$)*

We now study an environment where the good is intrinsically valued and the agents are vulnerable to deceptive practices. In this environment, the SO can benefit both from the agents' sales and from their mistakes.

We impose two mild assumptions in this section. First, note that due to their risk neutrality and their different beliefs, both the agents and the SO may benefit from scaling up the license fees and the commissions by $\gamma > 1$. Thus, when $n > n_2^\star$, the SO's potential profit is infinite. To guarantee that the SO's potential profit is finite (and, therefore, a profit-maximizing scheme exists), we bound the agents' ability to pay for a license. Specifically, we assume that the maximal amount that each agent can pay for a license is B and that B is large with respect to c and $\hat{c}$. Second, to simplify the analysis, we fix an arbitrary integer $T > 2$ and impose that $a_\tau = b_\tau = 0$ for every $\tau > T$. Finally, without loss of generality, we maintain the assumption that $\eta = q$.

We start the analysis with the following observation: when q is sufficiently small, profit-maximizing schemes charge a license fee and pay recruitment commissions just like the schemes discussed in Section 4.1. When q goes to zero, the SO's potential sales revenue goes to zero as well. However, the potential profit from the agents' mistakes does not vanish: agents are still willing to pay for distribution licenses given the "right" scheme and a budget B that is not too small. To maximize his expected profit, the SO takes advantage of this mistake, which as noted in the previous section, requires charging a license fee and paying for at least two levels of downline recruits.

Next, we establish that when the potential gains from sales are sufficiently large, schemes that pay recruitment commissions, charge license fees, or pay for more than two levels of downline sales are not profit-maximizing.

Proposition 2. *Fix $q > 0$. There exists a number $\hat{n}(q)$ such that if $n > \hat{n}(q)$, then the profit-maximizing scheme is identical to the profit-maximizing scheme in Theorem 1.*

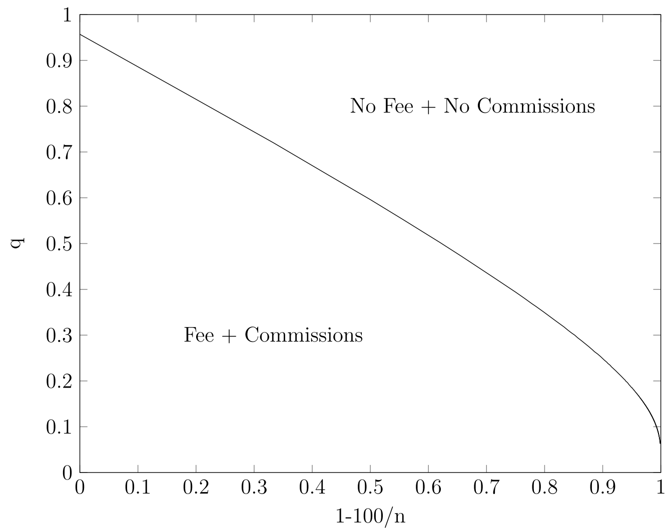

FIGURE 3. The components of the profit-maximizing scheme for $c = 0.25$, $B = 2.5$, $T = 3$, $\hat{c} = 0$, and $(n, q) \in \{100, 125, \dots, 100{,}000\} \times \{0.001, 0.002, \dots, 1\}$.

Charging a license fee enables the SO to make a profit from the agents' mistakes. However, as noted in Section 3, the license fee has an additional, indirect, negative effect on the SO's ability to make a profit from the agents' sales. A fee makes it more costly to become a distributor, which requires paying higher sales bonuses and recruitment commissions to attract prospective distributors. This effect becomes stronger when the potential profit from sales is large (i.e., when n and q are large). In such instances, the SO uses many distributors to increase the number of agents who purchase the good to benefit from the large demand. Thus, he has to pay multiple bonuses for every sale and multiple commissions for every recruitment such that raising the bonuses and commissions to compensate for the license fee becomes extremely costly and the SO is better off not charging a license fee. Recall that an IC scheme that does not charge a license fee cannot pay recruitment commissions either.

The above argument suggests that the SO faces a trade-off between making a profit from the agents' mistakes and making a profit from selling the good. When $n > n^\star$ and q are not too large, this trade-off is decided in favor of profiting from the agents' mistakes as the potential profit from sales is small. Thus, the SO charges a license fee, pays for recruitment, and may benefit from extending the scheme beyond two levels. For large values of n and q, this trade-off is decided in favor of profiting from the agents' sales, which as explained in the previous section, requires a 2-level scheme that pays only sales bonuses.

The main findings of this section are summarized in the following corollary and illustrated in Figure 3, which shows that for a given $n > n_2^\star$, the SO benefits from charging license fees and paying recruitment commissions if and only if the willingness to pay q is below some cutoff $q(n)$ that is decreasing in n.

Corollary 1. *There exists a number $\tilde{n}$ such that for every $n > \tilde{n}$:*

- *If q is sufficiently large, then in every profit-maximizing scheme it holds that $\mathbf{a} = \mathbf{b}_{-2} = \mathbf{0}$ and $\phi = 0$.*
- *If q is sufficiently small, then every profit-maximizing scheme charges $\phi > 0$ and pays at least two recruitment commissions. Moreover, it is possible that $a_\tau > 0$ for $\tau > 2$.*

We conclude that license fees and recruitment commissions are useful for the SO when the demand is low, but not when it is high. This result is related to Heidhues, Kőszegi, and Murooka's (2017) findings in the context of hidden add-on prices. They show that firms tend to benefit from this deceptive practice only in markets in which socially invaluable products (i.e., with lower social surplus than an alternative) are traded. The reason for this effect in their setting is that, when a product is socially valuable, firms may be able to increase market share and profits by drawing the consumers' attention to the hidden fees and undercutting prices.

4.3 *Public policy*

A natural way to shut down pyramids when $q = 0$ is to ban the *combination* of license fees and multiple recruitment commissions. However, as noted above, when $q > 0$ is small, this policy reduces the SO's potential profit. Thus, it may force the SO to shut down if there is a fixed cost for operating the scheme. For the following proposition, we assume that there exists such a cost and denote it by c^{SO}.[17] We say that the policy forces the SO to shut down if his potential profit under the policy is less than c^{SO}. Let $\mathcal{R}$ be the set of all schemes and $\mathcal{R}^{\mathrm{reg}}$ be the set of schemes in which $(\phi, \mathbf{a}, \mathbf{b}_{-2}) = (0, \mathbf{0}, \mathbf{0})$. The next proposition shows that if the operation is viable with rational agents (i.e., $\max_{R\in\mathcal{R}} \pi^{\mathrm{PBE}}(R) \geq c^{\mathrm{SO}}$), then the SO will not be forced to shut down due to the proposed policy regardless of whether agents are rational or not.

Proposition 3. *It holds that*

$$\max_{R\in\mathcal{R}^{\mathrm{reg}}} \pi^{\mathrm{ABEE}}(R) \geq \max_{R\in\mathcal{R}^{\mathrm{reg}}} \pi^{\mathrm{PBE}}(R) = \max_{R\in\mathcal{R}} \pi^{\mathrm{PBE}}(R).$$

Theorem 1 showed that when agents are fully rational there is a profit-maximizing scheme R in which $(\phi, \mathbf{a}, \mathbf{b}_{-2}) = (0, \mathbf{0}, \mathbf{0})$. A profile of strategies is part of a PBE of $\Gamma(R)$ if and only if it is also part of an ABEE of $\Gamma(R)$. To see why, note that in a PBE of $\Gamma(R)$, a distributor who meets an agent will try to recruit the latter if and only if he expects to recoup the cost $\hat{c}$ from the bonuses from the agent's sales. Moreover, an agent will purchase a license if and only if he expects to recoup the cost c from the bonuses from his own sales. The equivalence between the solution concepts then follows from the fact that agents' analogy-based expectations regarding nonstrategic aspects and, in particular, regarding the number of future sales of the good conditional on holding a licence are correct.

[17] Such a cost can be added to the baseline model without changing the paper's results.

In addition to its effect on the SO, the policy in Proposition 3 influences the agents. The policy may lead the SO to use a scheme that incentivizes a smaller number of agents to become distributors, which means that fewer agents get to consume the good.[18] Recall that under PBE too few agents become distributors in the profit-maximizing scheme relative to the social optimum, akin to the classic monopoly distortion. Under ABEE, whether there are too many or too few distributors relative to the social optimum is undetermined. Roughly speaking, for small values of n and q the number of distributors in the profit-maximizing scheme is greater than the socially optimal number of distributors, whereas for larger values of n and q the number of distributors in the profit-maximizing scheme is smaller than the socially optimal number of distributors. Thus, from a social perspective, regulatory policies that lead to a smaller number of distributors are beneficial in the former case but are detrimental in the latter case.

Note that a more moderate policy that would still shut down pyramids that are based on stationary reward schemes is to ban the *combination* of $\phi > 0$ and $\mathbf{a}_{-1} \neq \mathbf{0}$. This policy may be safer than the one suggested in Proposition 3 as, in reality, license fees and referral commissions are prevalent in many business models for legitimate reasons that are beyond the scope of our model, whereas the combination of $\phi > 0$ and $\mathbf{a}_{-1} \neq \mathbf{0}$ is perhaps not as prevalent outside the context of MLM.

5. Concluding remarks

Legitimate MLM and fraudulent pyramid scams are two widespread phenomena. We developed a model that promotes a better understanding of the compensation structures that underlie these phenomena and explains why MLM compensation structures appear to be correlated with scams. The paper shows that a company can make a profit even in instances in which participants' beliefs are statistically correct and its product has no intrinsic value. Sustaining such a scam requires the company to charge a license fee and pay for at least two levels of downline recruits. The paper's benchmark results show that companies with a good product that face rational agents find it detrimental to use these two tools or to extend the scheme beyond two levels. These results provide a potential explanation of why passive income from recruiting others is viewed as one of the hallmarks of pyramid scams (SEC (2013)).

We shall conclude by discussing several modifications of the behavioral model, the incentive constraint, and the network formation process.

Alternative behavioral models

The main impossibility argument in the paper relies on an assumption (which is embedded in ABEE) that agents have *statistically correct beliefs* rather than on other specific properties of the solution concept. For example, it would hold under any assumption

[18] In our model, due to the simple demand structure, consumers are left with no surplus from personal consumption, and so their welfare is unaffected by the policy. However, under a more realistic assumption about the agents' utility from personal consumption, the smaller number of distributors could negatively affect the consumers' welfare.

about the agents' analogy partitions. This argument can be made under any behavioral model (adapted to the baseline framework) in which agents' beliefs are statistically correct in the same sense as in an ABEE. Such models include the rational expectations model, the partially cursed equilibrium (Eyster and Rabin (2005)), and other models of causal misperception (see Spiegler (2016), for a modeling framework).

In the working paper version of this article (Antler (2021)), we adapt Brunnermeier and Parker's (2005) model of optimal expectations to the framework and revisit the results of Section 4.1. Brunnermeier and Parker's model yields motivated beliefs that are not statistically correct, and so the logic of the impossibility argument breaks down. While the results are not as sharp as under ABEE, it is shown that extending the number of levels of the scheme increases a scam's profits and can make the scam viable in cases where this is not viable with a smaller number of levels. Thus, even when 1-level schemes yield positive profits from a worthless good, the "optimal" pyramid scam includes a license fee and at least two levels of recruitment commissions.

In general, a variety of behavioral biases can support pyramid scams. However, not all biases support such scams: some biases might be orthogonal, or even reduce the profitability of scams. Any behavioral bias that can support a pyramid scam must be able to overcome the challenge that typically pyramid scams will collapse, and backward induction typically will unravel the pyramid from this point backwards, implying that pyramids never start. This suggests that a behavioral bias that supports a pyramid scam must lead the last agents to join the pyramid to not realize that they are the last to join. We presented one example of a model that leads to such an error: the ABEE in which agents do not grasp the time-contingent nature of the other agents' behavior. In this model and others, the bias that causes agents to overestimate their number of recruits will also cause them to overestimate the recruitment success of their recruits, and that of their recruits' recruits, and so on. As a result, the size of the mistake that can be exploited accumulates to a larger and more profitable sum as more levels are incorporated into the reward scheme. We conclude that the more levels the scheme has, the more profitable it is for the SO to exploit a given bias via a pyramid scam.

Stronger incentive compatibility

Throughout the paper, we assumed that the SO uses only IC schemes to prevent distributors from manipulating him by creating fictitious players. The incentive-compatibility constraint prevents these manipulations when the SO can verify the identity of any distributor who wishes to be paid (in practice, to be paid, MLM distributors are often required to identify themselves). An SO who cannot verify the distributors' identities may wish to use a reward scheme where $\sum_{\tau=1}^{n} a_\tau \leq \phi$ and $\sum_{\tau=1}^{n} b_\tau \leq q$ to prevent each distributor from creating a tree of fictitious recruits and collecting the commissions that all the nodes in the tree would be eligible for.

A natural question is whether the SO can sustain a pure pyramid scam under the stronger IC constraint $\sum_{\tau=1}^{n} a_\tau \leq \phi$ and $\sum_{\tau=1}^{n} b_\tau \leq q$. The answer depends on the social network in which the scheme operates: if all agents have, on average, more friends/successors in the graph than their friends/successors in the graph have, then it is impossible to sustain a pyramid scam under the stronger IC constraint. To gain intuition for this, recall that, in an ABEE, the marginal agent expects to recruit less than one distributor. If he expects that his successors will have fewer opportunities to recruit new members, then he must expect each of them to recruit fewer than one person and, as a result, to have fewer than one distributor in every level of his downline. Thus, if $\sum_{\tau=1}^{n} a_\tau \leq \phi$, the marginal agent cannot analogy-based expect to recoup the fee paid. In the random recursive tree model, which we used as our network formation model throughout the paper, this property holds as v_t is decreasing in t and, therefore, it is impossible for the SO to sustain a pure pyramid scam under the stronger IC constraint. As we illustrate in the working paper version (Antler (2021)), when the network is such that some agents have fewer successors than their successors, it can be possible to sustain a pure pyramid scam under the stronger IC constraint; however, this requires paying at least two recruitment commissions and charging a license fee.

Alternative social network models

The above discussion naturally leads to a discussion of the choice of the network formation model. Throughout the analysis, we used a simple network formation process that allowed us to succinctly describe the social network in which the SO operates. Underlying this process are the assumptions that (1) the horizon is finite, (2) only one new agent enters in each period, and (3) each agent is connected to only one agent upstream. However, our main insights do not depend on these assumptions.[19] In particular, the main arguments of Section 4.1 hinge on the agents' statistically correct beliefs and the fact that, on average, each distributor recruits less than one person, and these properties are orthogonal to the social network. We therefore expect our main insights to hold under any well-behaved social network model.

Appendix: Proofs

A.1 *Results of Section 3*

Proof of Proposition 1. We prove this result by backward induction. Consider a scheme R and a PBE of $\Gamma(R)$. Sequential rationality implies that agent i_n rejects every offer he receives, both on and off the equilibrium path. Let $t \in \{1, \dots, n-1\}$ and suppose that all agents reject every offer they receive (both on and off the path) after period t in our PBE. Agent i_t cannot expect to sell licenses. Since $q = 0$, incentive compatibility implies that $b_1 = 0$. Being sequentially rational, agent i_t rejects every offer he receives both on and off the equilibrium path. By induction, no agent purchases a license in our PBE and, since $q = 0$, $\pi^{\text{PBE}}(R) = 0$. □

[19] In the working paper version (Antler (2021)), we studied different models of social networks and showed that the main results of Section 4.1 hold, even under the stronger incentive constraint. In particular, we studied a network formation model with an infinite horizon, a network formation model in which multiple agents enter each period, and models of deterministic networks in which agents are connected to multiple agents upstream.

Proof of Lemma 1. In expectation, independently of the players' strategies and the events that took place up to period t, an agent who purchases a license in period t meets v_t agents who purchase the good. Because of this independence, in a PBE, both on and off the equilibrium path, all agents correctly expect that, conditional on purchasing a license, agent i_t will sell v_t units of the good.

Part 1. For sufficiency, note that selling v_t units yields a payoff of $v_t b_1 - c - \phi$ to agent i_t even if the latter does not sell licenses. For necessity, let t be the last period in which an agent buys a license in a PBE (on and off the equilibrium path). Agent i_t cannot expect to sell licenses in this PBE. Thus, he can expect a payoff of at most $v_t b_1 - c - \phi$ conditional on purchasing a license. Sequential rationality implies that $v_t b_1 - c - \phi \geq 0$. Since v_t is decreasing in t, it holds that $b_1 v_k \geq c + \phi$ for every $k \leq t$.

Part 2. For sufficiency, note that a distributor who recruits agent i_t earns a_1 for the recruitment and b_2 for every unit i_t sells. Thus, if $a_1 + b_2 v_t \geq \hat{c}$ the distributor finds it optimal to recruit i_t even if he does not expect i_t to sell licenses. For necessity, consider a PBE and let t be the last period in which both (i) $v_t b_1 \geq c + \phi$ and (ii) a distributor is supposed to make an offer to an agent (on and off the equilibrium path). Since $v_t b_1 \geq c + \phi$, agent i_t accepts the offer. Moreover, the distributor cannot expect i_t to recruit anyone and, therefore, the distributor believes that making an offer to i_t yields an expected payoff of $a_1 + b_2 v_t - \hat{c}$. Sequential rationality implies that $a_1 + b_2 v_t - \hat{c} \geq 0$. Since v_t is decreasing in t, $a_1 + b_2 v_k \geq \hat{c}$ holds for every $k \leq t$. □

Proof of Theorem 1. Let $R = (\phi, a_1, b_1, b_2)$ be an IC 2-level scheme and denote $k_1 = \sup\{t | b_1 v_t \geq \phi + c\}$. If $a_1 + b_2 v_1 < \hat{c}$, let $k_2 = 1$ and, otherwise, let $k_2 = \sup\{t | a_1 + b_2 v_t \geq \hat{c}$ and $v_t b_1 \geq \phi + c\}$. Step 1 will show that either (i) there exists an IC scheme $R' = (\phi', a_1', b_1', b_2') = (0, 0, c/v_{k_1}, \hat{c}/v_{k_2})$ such that $\pi^{\text{PBE}}(R') \geq \pi^{\text{PBE}}(R)$ or (ii) there exists an IC scheme $R'' = (\phi'', a_1'', b_1'', b_2'') = (0, 0, c/v_{k_1}, 0)$ such that $\pi^{\text{PBE}}(R'') \geq \pi^{\text{PBE}}(R)$. Moreover, Step 1 will show that the above inequalities are strict if $\phi > 0$ and R incentivizes more than one agent to purchase a license.

Step 1. Since v_t is decreasing in t, Lemma 1 implies that in every PBE of $\Gamma(R)$ and every PBE of $\Gamma(R')$ agents accept (resp., reject) every offer made up to (resp., after) period k_1, and every distributor recruits (resp., does not recruit) every agent he meets up to (resp., after) period k_2. Denote an arbitrary profile of strategies in which agents behave in this manner by σ. To show that $\pi^{\text{PBE}}(R') \geq \pi^{\text{PBE}}(R)$, we only need to show that the SO's expected payoff given σ is greater in $\Gamma(R')$ than in $\Gamma(R)$.

Given σ, the SO's revenue from sales is the same under both schemes. We now show that the expected net transfers (i.e., after subtracting the license fees) from the SO to the distributors when σ is played are greater in $\Gamma(R)$.

Consider the expected net transfers from the SO to a distributor i_t. If $d(\text{SO}, i_t) = 1$ (i.e., the SO recruits i_t), then under R, the SO collects ϕ from i_t and pays him b_1 for each of his sales. Under R', the SO pays i_t b_1' for each sale. Note that

$$-\phi + v_t b_1 \geq -\phi + v_t\left(\frac{c+\phi}{v_{k_1}}\right) \geq v_t \frac{c}{v_{k_1}} = v_t b_1', \tag{4}$$

with the second inequality being weak for $t = k_1$ and *strict* if $t < k_1$ and $\phi > 0$.

If $d(\mathrm{SO}, i_t) > 1$ (i.e., i_t is recruited by a distributor $j \in I$), then under R, the SO collects ϕ from i_t, pays a_1 to j, and for each of i_t's sales, the SO pays b_1 to i_t and b_2 to agent j. Under R', the SO pays b_1' to i_t and b_2' to agent j for each retail sale made by i_t. The SO's expected net transfers to the distributors are greater under R if

$$a_1 - \phi + v_t(b_1 + b_2) \geq a_1\left(1 - \frac{v_t}{v_{k_2}}\right) - \phi\left(1 - \frac{v_t}{v_{k_1}}\right) + \frac{v_t \hat{c}}{v_{k_2}} + \frac{v_t c}{v_{k_1}} \geq v_t\left(b_1' + b_2'\right). \tag{5}$$

Since R is IC, $a_1 \leq \phi$. As $t \leq k_2 \leq k_1$, (5) holds. We conclude that for every $t \leq k_1$ the expected net transfers based on i_t's recruitment and sales are weakly lower in R' under σ. Hence, $\pi^{\mathrm{PBE}}(R') \geq \pi^{\mathrm{PBE}}(R)$. Moreover, if $\phi > 0$ and the SO makes an offer in some period $t < k_1$, then $\pi^{\mathrm{PBE}}(R') > \pi^{\mathrm{PBE}}(R)$.

Note that $b_1 \leq q$ since R is IC and that $b_1' \leq b_1$. If $b_2' > q$, then in $\Gamma(R')$, the SO incurs a loss of $b_1' + b_2' - q$ whenever a sale is made by a distributor i such that $d(\mathrm{SO}, i) > 1$. The SO can earn more than $\pi^{\mathrm{PBE}}(R')$ by using a scheme R'' that is identical to R' except that $b_2'' = 0$ as distributors never recruit in $\Gamma(R'')$. Thus, if R' is not IC, then R'' is IC and $\pi^{\mathrm{PBE}}(R'') > \pi^{\mathrm{PBE}}(R)$.

Step 2. By Step 1, there exists a profit-maximizing scheme that takes the form of either (i) $(\phi, a_1, b_1, b_2) = (0, 0, c/v_{k_1}, \hat{c}/v_{k_2})$ where $k_1 \in \{1, \dots, n\}$ and $k_2 \in \{1, \dots, k_1\}$ or (ii) $(\phi, a_1, b_1, b_2) = (0, 0, c/v_{k_1}, 0)$ where $k_1 \in \{1, \dots, n\}$. In this class of schemes, we can identify schemes with the cutoffs k_1 and k_2 they induce (we set $k_2 = 0$ when referring to schemes of class (ii) and use $k_1 = k_2 = 0$ to represent the case where the SO refrains from selling licenses). Note that if a scheme (k_1, k_2) in this class is not IC, then $b_1 > q$ or $b_1 + b_2 > q$. If $b_1 > q$, then the SO incurs a net loss from sales made by distributors and so $(0, 0)$ yields a higher expected SO payoff. If $b_1 + b_2 > q$, then the SO incurs a net loss from sales of distributors he did not recruit directly and so $(k_1, 0)$ yields a higher expected SO payoff. Hence, the profit-maximizing scheme in this class is IC.

Consider a scheme (k_1, k_2) with $k_1 > 0$ and $b_1 + b_2 \leq q$. The SO's expected net payoff based on agent $i_{\max\{k_1,k_2\}}$'s sales is $qv_{\max\{k_1,k_2\}} - \hat{c} - c$. If $qv_{\max\{k_1,k_2\}} - \hat{c} - c \geq 0$, then the SO finds it optimal to make an offer in every period $t \leq k_1$ as the SO's expected net payoff based on agent i_t's sales is higher even without taking into account the possibility that i_t may recruit new members. If $qv_{\max\{k_1,k_2\}} - \hat{c} - c < 0$, then (k_1, k_2) is not profit-maximizing: the scheme $(k_1 - 1, \min\{k_2, k_1 - 1\})$ pays lower commissions and does not incentivize the recruitment of $i_{\max\{k_1,k_2\}}$. Hence, if (k_1, k_2) is a profit-maximizing scheme and $k_1 > 0$, then in a PBE of its induced game the SO makes an offer to every agent he meets.

Denote by $\Pi(k_1, k_2)$ the SO's payoff given that he makes an offer to every agent he meets and the agents' behavior is described by the cutoffs (k_1, k_2). Observe that

$$\begin{aligned}\Pi(k_1, k_2) = {} & qk_2(1 + v_{k_2}) + q\sum_{j=k_2+1}^{k_1} \frac{(1 + v_j)}{j} + qv_{k_1} \\ & - \sum_{j=1}^{k_2}\left(\frac{v_j}{j}\left(\frac{c}{v_{k_1}} + \frac{\hat{c}}{v_j}\right) + \frac{(j-1)v_j}{j}\left(\frac{c}{v_{k_1}} + \frac{\hat{c}}{v_{k_2}}\right)\right)\end{aligned}$$

$$-\sum_{j=k_2+1}^{k_1} \frac{v_j}{j}\left(\frac{c}{v_{k_1}}+\frac{\hat{c}}{v_j}\right), \tag{6}$$

where $\sum_{j=t+1}^{t} := 0$. We now study the cutoffs that maximize Π. These cutoffs yield the maximal expected payoff that the SO can obtain using an IC scheme.

If $q < (c+\hat{c})/v_1$, then $\Pi(0,0) > \Pi(k_1,k_2)$ for every $k_1 > 0$. Hence, the SO's profit is maximized by not selling licenses, and so all schemes are vacuously profit-maximizing.

If $q > (c+\hat{c})/v_2 + c/(2v_2^2)$, then $\Pi(2,2) > \max\{\Pi(2,0), \Pi(1,1), \Pi(1,0), \Pi(0,0)\}$, which implies that, in the profit-maximizing scheme among this class of schemes, $k_1 \geq k_2 \geq 2$. We can apply Step 1 to show that (i) any scheme R that induces a different PBE behavior is not profit-maximizing, (ii) any scheme R that charges $\phi > 0$ (and, by incentive compatibility, any scheme that pays $a_1 > 0$) is not profit-maximizing. Since at the optimum agent i_{k_1} is indifferent between purchasing a license and not doing so, and a distributor who meets agent i_2 is indifferent between recruiting i_2 and not doing so, we obtain that $b_1 v_{k_1} = c$ and $b_2 v_{k_2} = \hat{c}$. Hence, $(\phi, a_1, b_1, b_2) = (0, 0, c/v_{k_1}, \hat{c}/v_{k_2})$ is the unique profit-maximizing scheme.[20]

If $q \in ((c+\hat{c})/v_1, (c+\hat{c})/v_2 + c/(2v_2^2))$, then $(k_1,k_2) = (1,0)$ is the unique maximizer of (6).[21] We can apply Step 1 to show that any scheme that induces a different PBE behavior is not profit-maximizing. All schemes that incentivize the first entrant to purchase a license and leave him with no surplus are profit-maximizing. Hence, $b_1 = (c+\phi)/v_1$ in this case. □

A.2 *Results of Section 4*

Proof of Theorem 2. Consider an IC 1-level scheme R. If $a_1 \leq \hat{c}$, then conditional on purchasing a license, an agent gets an expected payoff of at most $-c-\phi < 0$. Hence, no agent purchases a license in an ABEE of $\Gamma(R)$.

Suppose that $a_1 > \hat{c}$ and consider an ABEE of $\Gamma(R)$ in which the SO makes offers. Clearly, agent i_n rejects every offer in this ABEE. Let $t \in \{1,\dots,n-1\}$ and suppose that agents reject every offer they receive in periods $t+1,\dots,n$. Since $a_1 \geq \hat{c}$, every agent who holds a license at the end of period t makes, in expectation, v_t offers in periods $t+1,\dots,n$. Thus, for every offer that is accepted in periods $1,\dots,t$ there are, in expectation, v_t rejected offers in periods $t+1,\dots,n$, and so the proportion of accepted offers, α_1, cannot exceed $1/(1+v_t)$. Hence, conditional on accepting an offer, agent i_t falsely expects a payoff of $\alpha_1 v_t(a_1-\hat{c}) - c - \phi < a_1 - \phi$. Since R is IC, $a_1 \leq \phi$, and so i_t rejects every offer he receives in our ABEE. We can conclude that no agent purchases a license in our ABEE. Hence, $\pi^{\text{ABEE}}(R) = 0$. □

Proof of Theorem 3. The proof consists of four parts. Part 1 shows that if $\pi^{\text{ABEE}}(R) > 0$ and R is a 2-level scheme, then $a_1 \geq \hat{c}$. Part 2 shows that if there exists a 2-level scheme R such that $\pi^{\text{ABEE}}(R) > 0$ when there are n agents, then there exists an IC 2-level scheme

[20] We ignore the nongeneric case in which there exist multiple pairs (k_1,k_2) that maximize (6).

[21] To see this, note that, in this parameter range, $\Pi(1,0) > \Pi(2,2)$, $\Pi(k,k) > \Pi(k+1,k+1)$ for $k > 1$, and $\Pi(k_1,k_2) > \Pi(k_1+1,k_2)$ for $k \geq 1$.

R' such that $\pi^{\text{ABEE}}(R') > 0$ when there are $n' > n$ agents. Part 3 shows that $a_1 + a_2 > \phi$ is a necessary condition for a 2-level scheme to be profitable. Part 4 shows that a profitable 2-level scheme exists if n is sufficiently large.

Part 1. Assume to the contrary that agents purchase licenses in an ABEE of an IC 2-level scheme R in which $a_1 < \hat{c}$. A distributor who recruits agent i_t expects to increase his payoff by $a_1 - \hat{c} + \alpha_1\alpha_2 v_t a_2$. Since v_t is decreasing in t, there is a cutoff k_2 such that distributors make offers to every agent they meet up to period k_2 and make no offers afterward. Moreover, no agent ever purchases a license in any period $t \geq k_2$ as he knows that he will refrain from making offers. It follows that an agent who purchases a license in period $t < k_2$ (analogy-based) expects a payoff of

$$\sum_{j=t+1}^{k_2} \frac{1}{j}\alpha_1\big[(a_1 - \hat{c}) + \alpha_1\alpha_2 v_j a_2\big] - c - \phi \leq \alpha_1\alpha_2 v_t(v_t - v_{k_2})\alpha_1 a_2 - c - \phi. \tag{7}$$

Consider the last period t in which an agent accepts an offer in our ABEE. Each offer accepted in periods $1, \dots, t$ leads in expectation to (i) $v_t - v_{k_2}$ rejected offers after period t and (ii) v_t opportunities to make offers where v_{k_2} of them are not made. Thus, $\alpha_1 \leq 1/(1 + v_t - v_{k_2})$ and $\alpha_2 \leq (1 + v_t - v_{k_2})/(1 + v_t)$. Hence, $\alpha_1^2\alpha_2 v_t(v_t - v_{k_2}) < 1$. As R is IC, $a_2 \leq \phi$. We conclude that (7) is strictly negative, which violates the optimality of purchasing a license in this ABEE.

Part 2. Consider an IC 2-level scheme R such that $\pi^{\text{ABEE}}(R) > 0$. By Part 1, $a_1 \geq \hat{c}$ such that, in an ABEE, every distributor makes an offer to every agent he meets. In an ABEE that maximizes the SO's expected payoff, there is a period k such that i_t accepts (resp., rejects) every offer he receives if $t \leq k$ (resp., if $t > k$) and the SO makes no offers after period k (as such offers are rejected and lower the agents' expectations). In this ABEE, $\alpha_1 = 1/(1 + v_k)$ and $\alpha_2 = 1$. Hence, agent i_k falsely expects a payoff of

$$\frac{v_k}{1 + v_k}(a_1 - \hat{c}) + \frac{\displaystyle\sum_{j=k+1}^{n-1} \frac{v_j}{j}}{(1 + v_k)^2} a_2 - c - \phi \geq 0, \tag{8}$$

where $v_k/(1 + v_k)$ is the number of agents he analogy-based expects to recruit and $(\sum_{j=k+1}^{n-1} v_j/j)/((1 + v_k)^2)$ is the number of agents he analogy-based expects his recruits to recruit.

Note that (8) is increasing in n. Thus, for any $n' > n$, there is a scheme $R^{n'}$ identical to R except that $\phi^{n'} > \phi$ such that the profile of strategies described in the above paragraph is part of an ABEE of $\Gamma(R^{n'})$. Hence, if $\pi^{\text{ABEE}}(R) > 0$ when there are n agents, then $\pi^{\text{ABEE}}(R^{n'}) > 0$ when there are n' agents.

Part 3. To see that $a_1 + a_2 > \phi$ is a necessary condition for a 2-level scheme to be profitable, fix k and note that when n goes to infinity, (8) converges monotonically to $a_1 - \hat{c} - 0.5a_2 - c - \phi$. If $a_1 + a_2 \leq \phi$, then i_k finds it suboptimal to purchase a license, and so an ABEE in which agents purchase licenses cannot exist.

Part 4. Consider a profile of strategies σ in which agent i_1 accepts an offer if he receives one and all other agents reject every offer they receive, the SO makes an offer only in period 1, and conditional on purchasing a license, every agent makes an offer to every agent whom he meets. The SO's expected payoff is $\phi - \hat{c}$ under σ. Agent i_1's analogy-based expected payoff in the ABEE that corresponds to σ is given in the LHS of (8) for $k = 1$. When n goes to infinity, (8) goes to $a_1 - \hat{c} + 0.5a_2 - c - \phi$. Thus, for a sufficiently large n, we can choose ϕ, $a_1 \in (\hat{c}, \phi]$, and $a_2 \leq \phi$ such that (8) holds in equality and σ is part of an ABEE in which the SO's payoff is strictly positive. □

Proof of Theorem 4. Consider an IC 2-level scheme R. In Part 2 of the proof of Theorem 3, we showed that if $\pi^{\text{ABEE}}(R) > 0$, then (8) must hold for some $k \leq n$. For $a_1 \leq \phi$ and $a_2 \leq \phi$, the LHS of (8) is smaller than

$$\frac{v_k}{1+v_k}\phi + \frac{\sum_{j=k+1}^{n} \frac{v_j}{j}}{(1+v_k)^2}\phi - \phi - c,$$

which is strictly negative if $\phi \geq 0$, $n \leq 25$, and $k < n$.

We now show that for $n \geq 25$ there exists an IC 3-level scheme R such that $\pi^{\text{ABEE}}(R) > 0$. Fix a profile σ as described in Part 4 of the proof of Theorem 3 and recall that σ is consistent with $\alpha_1 = 1/(1+v_1)$ and $\alpha_2 = 1$ and induces an expected payoff of $\phi - \hat{c}$ to the SO. Consider a 3-level scheme R in which $a_1 = a_2 = a_3 = x\phi$. Under σ, the first entrant obtains an analogy-based expected payoff of

$$\alpha_1 \sum_{j=2}^{n} \frac{(x\phi - \hat{c})}{j} + \alpha_1^2\alpha_2 \sum_{j=2}^{n-1} \sum_{j'=j+1}^{n} \frac{x\phi}{jj'} + \alpha_1^3\alpha_2^2 \sum_{j=2}^{n-2} \sum_{j'=j+1}^{n-1} \sum_{j''=j'+1}^{n} \frac{x\phi}{jj'j''} - \phi - c \tag{9}$$

and, conditional on purchasing a license, every agent who enters the game after period 1 obtains less than (9). For $n \geq 25$, it is possible to choose a large ϕ and an $x < 1$ such that (9) equals 0, in which case, the first entrant finds it optimal to purchase a license while all other agents find it optimal not to purchase a license. Moreover, we can choose x such that $\phi x > \hat{c}$ and the distributors find it optimal to make an offer to every agent they meet. Hence, given the chosen ϕ and x, σ is part of an ABEE of $\Gamma(R)$ in which the SO makes a strictly positive expected payoff. □

Proof of Proposition 2. Let $(R^n)_{n=1}^{\infty}$ be a sequence of IC schemes such that each R^n is profit-maximizing when there are n agents. For each $n \in \mathbb{N}$, let (σ^n, α^n) be an ABEE of $\Gamma(R^n)$ that induces an expected profit of $\pi^{\text{ABEE}}(R^n)$, where $\alpha^n = (\alpha_1^n, \alpha_2^n)$. We use k_1^n to denote the last period in which the agents accept offers and k_2^n to denote the last period in which distributors make offers in σ^n.

Step 1: *Lower bounds.* Lemmata 2 and 3 show that there exists a number $\gamma > 0$ and an integer n' such that $\pi^{\text{ABEE}}(R^n), k_1^n, k_2^n > \gamma n$ for every $n > n'$.

Lemma 2. *Fix $\gamma' > 0$ such that $q\log(1/\gamma')(1 - \frac{c+\hat{c}}{(\log(1/\gamma')-1)}) - \hat{c} > \gamma'$. There exists a number $n(\gamma')$ such that for every $n > n(\gamma')$ it holds that $\pi^{\text{ABEE}}(R^n) \geq \gamma'^2 n$.*

Proof. Denote $k = \lceil \gamma' n \rceil$. Let R be a 1-level scheme in which $\phi = \hat{c}$, $a_1 = \hat{c}$, and $b_1 = (c + \hat{c})/v_k$. Consider a profile σ in which each agent accepts (resp., rejects) every offer up to (resp., after) period k and each distributor makes an offer to every agent he meets. It is easy to verify that σ is part of an ABEE of $\Gamma(R)$. For a sufficiently large n, it holds that $1 + v_k > \log(1/\gamma')$ and $b_1 < (c + \hat{c})/(\log(1/\gamma') - 1)$. The SO's expected payoff under σ is greater than $k(q[1 + v_k](1 - b_1) - \hat{c})$. Hence, for a sufficiently large n, it holds that

$$\pi^{\text{ABEE}}(R^n) \geq \pi^{\text{ABEE}}(R) \geq \gamma' n\left(q \log(1/\gamma')\left(1 - \frac{c + \hat{c}}{(\log(1/\gamma') - 1)}\right) - \hat{c}\right) \geq \gamma'^2 n. \tag{10}$$

$\square$

Lemma 3. *There exist $\bar{\gamma} \in (0, 1)$ and $n(\bar{\gamma}) \in \mathbb{N}$ such that $\min\{k_1^n, k_2^n\} > \bar{\gamma} n$ for every $n > n(\bar{\gamma})$.*

Proof. The SO obtains a revenue of at most $q + B$ from every distributor and q from each agent who meets a distributor but does not purchase a license. Hence, in an ABEE in which $k_2 \geq k_1$, the SO obtains an expected revenue of no more than

$$k_1^n(q + B + qv_{k_1^n}) + qv_{k_1^n} < k_1^n\left(2q + B + q\log\left(\frac{n}{k_1^n}\right)\right) + q(\log(n) + 1).$$

In an ABEE in which $k_2 < k_1$, the SO obtains an expected revenue of no more than

$$k_2^n(q+B+qv_{k_2^n}) + \sum_{j=k_2^n+1}^{k_1} \frac{qv_j + B}{j} + qv_{k_2^n} < k_2^n\left(2q+B+q\log\left(\frac{n}{k_2^n}\right)\right) + (q+B)(\log(n)+1)^2.$$

It is now possible to find a small $\bar{\gamma} > 0$ and $n(\bar{\gamma})$ such that if $n > n(\bar{\gamma})$ and $\min\{k_1^n, k_2^n\} \leq \bar{\gamma} n$, then the linear lower bound on the SO's expected profit in (10) is greater than both of the above upper bounds. $\square$

Step 2: *The dual problem.* The profit-maximizing scheme must minimize the SO's expected cost among the class of IC schemes that charge ϕ^n and in which σ^n is part of an ABEE of their induced game. This problem can be written as

$$\begin{aligned} &\min_{a_1,\ldots,a_T,b_1,\ldots,b_T} \sum_{\tau=1}^{T}\big(a_\tau \kappa(a_\tau) + b_\tau \kappa(b_\tau)\big) \\ \text{s.t. } (i_{k_1}) \quad & \sum_{\tau=1}^{T}\big[a_\tau w(a_\tau) + b_\tau w(b_\tau)\big] - \hat{c} w(a_1) \geq c + \phi^n \\ (i_{k_2}) \quad & \sum_{\tau=1}^{T}\big[a_\tau \hat{w}(a_\tau) + b_\tau \hat{w}(b_\tau)\big] \geq \hat{c} \\ (\text{IC}) \quad & a_1, a_2, \ldots \leq \phi^n \quad \text{and} \quad b_1, b_2, \ldots \leq q, \end{aligned} \tag{11}$$

where (i) $w(z)$ is i_{k_1}'s willingness to pay for a license due to $z \in \{a_1, b_1, \ldots, a_T, b_T\}$, (ii) $\hat{w}(z)$ is the corresponding distributors' benefit from recruiting agent i_{k_2} due to z, and (iii) $\kappa(z)$ is the SO's cost of using z (i.e., the expected number of times z will be paid).

The SO's costs. The SO pays a_τ for the recruitment of every distributor j such that $d(\mathrm{SO}, j) > \tau$. The expected number of such distributors is

$$\kappa(a_\tau) = \sum_{G' \in \mathcal{G}^n} \sum_{j \in G'} \frac{\mathbb{1}\big(d(\mathrm{SO}, j) > \tau\big)}{\min\{k_1^n, k_2^n\}!}, \tag{12}$$

where $\mathcal{G}^n$ is the set of rooted trees with $\min\{k_1^n + 1, k_2^n + 1\}$ nodes, and $\mathbb{1}(d(\mathrm{SO}, j) > \tau) \in \{0, 1\}$ is an indicator that equals 1 if and only if $d(\mathrm{SO}, j) > \tau$. The SO pays b_τ for every sale made by a distributor j such that $d(\mathrm{SO}, j) \geq \tau$. Thus, (i) $\kappa(b_\tau) = \kappa(a_\tau) + \kappa(a_{\tau-1}) v_{\min\{k_1^n, k_2^n\}}$ for $\tau > 1$, (ii) $\kappa(b_1) = k_1^n v_{k_1^n} + \kappa(a_1)$ if $k_2^n \geq k_1^n$, and (iii) $\kappa(b_1) = k_2^n v_{k_2^n} + \kappa(a_1) + \sum_{j=k_2^n+1}^{k_1^n} v_j/j$ otherwise.

Constraint i_{k_1}. The LHS is agent $i_{k_1^n}$'s willingness to pay for a license. If $k_1^n \geq k_2^n$, then $w(b_1) = q v_{k_1^n}$ and $w(z) = 0$ for any other commission or bonus z as $i_{k_1^n}$ does not expect to make offers in equilibrium. If $k_1^n < k_2^n$, then for $\tau > 1$, it holds that $w(a_\tau) = (\alpha_1^n)^\tau (\alpha_2^n)^{\tau-1} \sum_{j=k_1^n+1}^{k_2^n} l_{j,\tau-1}/j$ and $w(b_\tau) = (\alpha_1^n)^{\tau-1} (\alpha_2^n)^{\tau-2} \sum_{j=k_1^n+1}^{k_2^n} l_{j,\tau-1}/j$, where $l_{j,\tau}$ is the expected number of agents in the τth level of the subtree of G rooted at the jth entrant (an agent j' is in the τth level of that subtree if $d(j, j') = \tau$). For $\tau = 1$, $w(b_1) = v_{k_1}$ and $w(a_1) = \alpha_1^n \sum_{j=k_1^n+1}^{k_2^n} 1/j$. Note that $\hat{c} w(a_1)$ is $i_{k_1^n}$'s expected cost of training new recruits.

Constraint i_{k_2}. The LHS is the expected reward of a distributor who recruits i_{k_2}. Clearly, $\hat{w}(b_1) = 0$ and $\hat{w}(a_1) = 1$. If $\tau > 1$, then $\hat{w}(a_\tau) = (\alpha_1 \alpha_2)^{\tau-1} l_{k_2^n, \tau-1}$ and $\hat{w}(b_\tau) = (\alpha_1 \alpha_2)^{\tau-2} l_{k_2^n, \tau-1}$.

Step 3: *For large n, the profit-maximizing scheme is a* 2*-level scheme.* We present results on random trees that will show that if n is sufficiently large, then for any $z \in \{a_1, b_2\}$ and $z' \in \{a_3, b_3, \ldots, a_T, b_T\}$, it holds that $w(z)/\kappa(z) \geq w(z')/\kappa(z')$ and $\hat{w}(z)/\kappa(z) \geq \hat{w}(z')/\kappa(z')$ with at least one strict inequality. Because of the linearity of (11) in w, $\hat{w}$, and κ, it follows that if $z' > 0$ and $z' \in \{a_3, b_3, \ldots, a_T, b_T\}$, then $a_1 = \phi$ and $b_2 = q$. In this case, the SO's expected profit cannot exceed $\sum_{t=1}^n (B + q)/t + \sum_{t=1}^{n-1} \sum_{j=t+1}^n 1/(tj)$, as he earns (at most) $q + B$ from every agent he meets and q from every agent who meets these agents. The latter expression is smaller than γn for a sufficiently large n, in contradiction to the result obtained in Step 1. Thus, to prove this step we only need to establish the inequalities above. We now calculate the *"cost-benefit" ratios* w/κ and $\hat{w}/\kappa$.

Consider the following simple technical result.

Lemma 4. *For every $j \in \{1, \ldots, n-1\}$ and $\tau \geq 1$, it holds that $v_j l_{j,\tau} \geq 2 l_{j,\tau+1}$.*

Proof. Note that if $j + \tau > n$, then $l_{j,\tau} = 0$. Observe that $v_j l_{j,1} = 2 l_{j,2} + \sum_{z=j+1}^n 1/(z^2)$ if $j \leq n-1$. We prove the lemma by induction on the size of τ. We assume that $v_j l_{j,\tau-1} \geq 2 l_{j,\tau}$ and show that it implies that $v_j l_{j,\tau} \geq 2 l_{j,\tau+1}$. We can write the latter inequality as

$$v_j \left(\frac{l_{j+1,\tau-1}}{j+1} + \frac{l_{j+2,\tau-1}}{j+2} + \cdots \right) \geq 2 \left(\frac{l_{j+1,\tau}}{j+1} + \frac{l_{j+2,\tau}}{j+2} + \cdots \right). \tag{13}$$

We can combine the induction hypothesis with the fact that v_j is weakly decreasing in j to see that (13) holds. $\square$

Consider the SO's expected cost $\kappa(a_\tau)$ that is given in (12) and note that it is essentially the expected number of the first $x = \min\{k_1^n, k_2^n\}$ entrants whose distance from the SO is greater than τ. Next, we show that when x goes to infinity, the share of the first x entrants whose distance from the SO is greater than τ goes to 1.

LEMMA 5. *For every $\tau \in \mathbb{N}$, it holds that*

$$\lim_{n\to\infty} \sum_{G' \in \mathcal{G}^n} \sum_{j \in G} \frac{\mathbb{1}\big(d(\mathrm{SO}, j) > \tau\big)}{x x!} = 1. \tag{14}$$

PROOF. The tree G is a uniform random recursive tree rooted at the SO, and the distance $d(\mathrm{SO}, i_t)$ corresponds to the *insertion depth* of the $(t+1)$th node. Theorem 1 in Mahmoud (1991) establishes that the normalized insertion depth $M_i^\star = (M_i - \log(i))/\sqrt{\log(i)}$ has the limiting distribution $\mathcal{N}(0, 1)$, that is, the standard normal distribution. Thus, the proportion of nodes inserted at a distance greater than τ from the root on the LHS of (14) goes to 1 when the size of the random tree, x, goes to infinity (to obtain the LHS of (14), note that the number of trees of size $x+1$ is $x!$). □

By Lemma 3, $\min\{k_1^n, k_2^n\} > \overline{\gamma} n$ for a sufficiently large n. Thus, Lemma 5 implies that $\lim_{n\to\infty} \kappa(a_\tau)/\kappa(a_{\tau+1}) = 1$ and $\lim_{n\to\infty} \kappa(b_\tau)/\kappa(b_{\tau+1}) = 1$ for every $\tau \in \{1, \dots, T-1\}$.

We split the analysis into three cases: (1) $n = k_2^n > k_1^n$, (2) $n > k_2^n > k_1^n$, and (3) $k_2^n \le k_1^n$.

Case 1. The analogy-based expectations in this case are $\alpha_1^n = 1/(1 + v_{k_1^n})$ and $\alpha_2^n = 1$. Thus, $w(a_\tau)/w(b_\tau) = 1/(1 + v_{k_1^n})$. Recall that $\kappa(b_\tau) = \kappa(a_\tau) + \kappa(a_{\tau-1}) v_{k_1^n}$ for $\tau > 1$ and $\kappa(b_1) = \kappa(a_1) + k_1^n v_{k_1^n}$. By definition, $k_1^n > \kappa(a_1) > \cdots > \kappa(a_T)$. Thus, it holds that $w(a_\tau)/\kappa(a_\tau) > w(b_\tau)/\kappa(b_\tau)$. Moreover, by Lemma 5, $\kappa(a_\tau)/k_1^n$ goes to 1 as n goes to infinity for every $\tau \in \{1, \dots, T\}$. By Lemma 4, it holds that $w(a_\tau) \ge 2w(a_{\tau+1})$ and $w(b_\tau) \ge 2w(b_{\tau+1})$. We can conclude that, for a sufficiently large n, it holds that

$$\frac{w(a_1)}{\kappa(a_1)} > \frac{w(b_1)}{\kappa(b_1)} > \frac{w(a_2)}{\kappa(a_2)} > \frac{w(b_2)}{\kappa(b_2)} > \frac{w(z)}{\kappa(z)} \tag{15}$$

for every $z \in \{a_3, \dots, a_T, b_3, \dots, b_T\}$. Note that $\hat{w}(a_1) = 1 > 0 = \hat{w}(z)$ for every $z \in \{a_2, \dots, a_T, b_1, \dots, b_T\}$.

Case 2. The analogy-based expectations in this case are $\alpha_1 = 1/(1 + v_{k_1^n} - v_{k_2^n})$ and

$$\alpha_2^n = \frac{k_1^n - \displaystyle\sum_{t=1}^{k_1^n} \frac{1}{t} + k_1^n (v_{k_1^n} - v_{k_2^n})}{k_1^n - \displaystyle\sum_{t=1}^{k_1^n} \frac{1}{t} + k_1^n v_{k_1^n}},$$

where $\sum_{t=1}^{k_1^n} 1/t$ represents the SO's expected number of offers in periods $t = 1, \dots, k_1^n$. Hence, $\alpha_1^n \alpha_2^n \le 1/(1 + v_{k_1^n})$. Since $\sum_{t=1}^{k_1^n} 1/(t k_1^n)$ goes to 0 as n goes to infinity, for a sufficiently large n, $\alpha_1^n \alpha_2^n$ is arbitrarily close to $1/(1 + v_{k_1^n})$. Note that $w(a_\tau)/(\alpha_1^n \alpha_2^n) = w(b_\tau)$ for $\tau > 1$. Thus, for a sufficiently large n, $w(a_\tau)/\kappa(a_\tau)$ is arbitrarily close to $w(b_\tau)/\kappa(b_\tau)$

for $\tau > 1$. As above, we can use Lemma 4 to show that $w(a_\tau) \geq 2w(a_{\tau+1})$ for $\tau \geq 1$ and that $w(b_\tau) \geq 2w(b_{\tau+1})$ for $\tau > 1$. We can conclude that for a sufficiently large n, it holds that

$$\frac{w(a_1)}{\kappa(a_1)} > \frac{w(b_2)}{\kappa(b_2)} > \frac{w(z)}{\kappa(z)} \tag{16}$$

for every $z \in \{a_3, \ldots, a_T, b_3, \ldots, b_T\}$. We can repeat the above exercise for the ratio $\hat{w}/\kappa$ (the proof is similar to the one for the ratio w/κ and, therefore, it is omitted).

Case 3. The analogy-based expectations in this case are $\alpha_1^n = 1$ and

$$\alpha_2^n = \frac{k_2^n - \displaystyle\sum_{t=1}^{k_2^n} \frac{1}{t}}{k_2^n - \displaystyle\sum_{t=1}^{k_2^n} \frac{1}{t} + k_2^n v_{k_2^n} + \displaystyle\sum_{j=k_2^n+1}^{k_1^n} \frac{v_j}{j}}.$$

It is easy to see that $w(b_1) > 0$ and $w(z) = 0$ for every $z \in \{a_1, \ldots, a_T, b_2, \ldots, b_T\}$, as agent $i_{k_1^n}$ does not expect to sell licenses. It is left to consider the ratio $\hat{w}/\kappa$. Clearly, $\hat{w}(a_1) = 1 > v_{k_2^n}/(1 + v_{k_2^n}) \geq \alpha_1^n \alpha_2^n v_{k_2^n} = \hat{w}(a_2)$. Since $\alpha_1^n \alpha_2^n < 1/(1 + v_{k_2^n})$, Lemma 4 implies that $\hat{w}(a_\tau) \geq 2\hat{w}(a_{\tau+1})$ for every $\tau > 1$. Similarly, Lemma 4 implies that $\hat{w}(b_\tau) \geq 2\hat{w}(b_{\tau+1})$ for every $\tau > 1$. Furthermore, for a sufficiently large n, α_2^n is arbitrarily close to $1/(1 + v_{k_2^n})$. This implies that, for a sufficiently large n, $\hat{w}(a_\tau)/\kappa(a_\tau)$ is arbitrarily close to $\hat{w}(b_\tau)/\kappa(b_\tau)$ for any $\tau > 1$. We can conclude that, for a sufficiently large n, it holds that

$$\frac{\hat{w}(a_1)}{\kappa(a_1)} > \frac{\hat{w}(a_2)}{\kappa(a_2)} > \frac{\hat{w}(z)}{\kappa(z)} \quad \text{and} \quad \frac{\hat{w}(b_2)}{\kappa(b_2)} > \frac{\hat{w}(z)}{\kappa(z)} \tag{17}$$

for every $z \in \{a_3, \ldots, a_T, b_3, \ldots, b_T\}$.

Step 4: *For a sufficiently large n, it holds that $\phi = 0$.* Assume to the contrary that R^n charges $\phi > 0$. There are two cases to consider: (1) R^n pays $a_1 \geq c$ and (2) R^n pays $a_1 < c$. Consider case (1) and note that $a_1 \geq c$ implies that a distributor who meets an agent finds it optimal to make an offer to the latter. Thus, $k_2^n = n$. As a result, $\hat{w}(a_2) = \hat{w}(b_2) = \hat{w}(b_1) = 0 < \hat{w}(a_1) = 1$. By inequality (15), if $a_2 > 0$ or $b_2 > 0$, then $a_1 = \phi$ and $b_1 = q$, which implies that the SO's expected profit cannot exceed $\sum_{t=1}^{n}(B+q)/t$ (i.e., the expected revenue from his sales and recruitments). For a sufficiently large n, the latter expression is smaller than the lower bound γn established in Step 1. It follows that if R^n is profit-maximizing, then $a_2 = b_2 = 0$.

At the optimum, agent $i_{k_1^n}$ must be indifferent whether to purchase a license as, otherwise, it would be possible to reduce b_1 without changing the agents' equilibrium behavior. Thus, $b_1 v_{k_1^n} + (a_1 - \hat{c})w(a_1) = \phi + c$, or $b_1 = (\phi + c - (a_1 - \hat{c})w(a_1))/v_{k_1^n}$.

We now show that R^n is not profit-maximizing by introducing a 2-level scheme R' such that $\pi^{\text{ABEE}}(R') > \pi^{\text{ABEE}}(R^n)$. Let $R' = (\phi', a_1', a_2', b_1', b_2') = (0, 0, 0, c/v_{k_1^n}, \hat{c}/v_{k_1^n})$. There exists an ABEE of $\Gamma(R)$ in which agents (resp., distributors) accept (resp., make) offers in periods $1, \ldots, k_1^n$ and reject (resp., do not make) offers in periods $k_1^n + 1, \ldots, n$. The SO's revenue from the agents' sales in this ABEE is identical to his revenue from the agents' sales in the ABEE of $\Gamma(R^n)$ as the same agents purchase licenses in both ABEEs. We now show that the expected net transfers (i.e., including the license fee) from the SO to the agents under R^n are higher than under R'.

In the transition from R^n to R', the SO's profit is lowered by ϕk_1^n since there is no fee under R'. The change from b_1 to b_1' increases the SO's expected payoff by $(\kappa(b_1)(\phi - (a_1 - \hat{c})w(a_1)))/v_{k_1^n}$, the reduction in a_1 increases the SO's expected payoff by $a_1\kappa(a_1)$, and the addition to b_2 decreases the SO's expected payoff by $\kappa(b_2)\hat{c}/v_{k_1^n}$. Overall, the SO's expected payoff increases in the transition from R^n to R' if

$$-\phi k_1^n + \frac{\phi - (a_1 - \hat{c})w(a_1)}{v_{k_1^n}}\kappa(b_1) + a_1\kappa(a_1) - \frac{\hat{c}}{v_{k_1^n}}\kappa(b_2) > 0.$$

Plugging $\kappa(b_1)$, $\kappa(b_2)$, and $w(a_1)$ yields

$$\frac{\phi\kappa(a_1)}{v_{k_1^n}} - (a_1 - \hat{c})\frac{\left(\kappa(a_1) + v_{k_1^n}k_1^n\right)}{1 + v_{k_1^n}} + (a_1 - \hat{c})\kappa(a_1) - \frac{\hat{c}\kappa(a_2)}{v_{k_1^n}} > 0.$$

By incentive compatibility, $\phi \geq a_1$ and, by assumption, $a_1 \geq \hat{c}$. By definition, $\kappa(a_1) \geq \kappa(a_2)$. Thus, the above inequality holds if

$$\frac{v_{k_1^n}\left(2\kappa(a_1) - k_1^n - \kappa(a_2)\right)}{(1 + v_{k_1^n})} > 0.$$

Note that $k_1^n - \kappa(a_1)$ is equal to the k_1^nth harmonic number, as it is equal to the expected number of nodes in the first level of a random recursive tree. Similarly, $\kappa(a_1) - \kappa(a_2)$ is equal to the expected number of nodes in the second level of a random recursive tree of size k_1^n, which is $\sum_{j=1}^{k_1^n-1}\sum_{j'=j+1}^{k_1^n} 1/(jj')$. Thus, for a sufficiently large k_1^n, it holds that $2\kappa(a_1) - k_1^n - \kappa(a_2) > 0$. By Lemma 3, $k_1^n > \overline{\gamma}n$ for sufficiently large n, and so the above inequality holds. Hence, R^n is not profit-maximizing. We conclude that, for a sufficiently large n, a scheme is profit-maximizing only if it charges $\phi = 0$. Incentive compatibility implies that $\mathbf{a} = \mathbf{0}$ at the optimum as well.

The proof for case 2 ($a_1 < \hat{c}$) is similar (for any scheme with $\phi > 0$, the transition to the scheme R' increases the SO's expected payoff) and, therefore, it is omitted.

Step 5: *The scheme R^n is identical to the profit-maximizing scheme in Theorem 1*. By the previous steps, we can focus on 2-level schemes in which $(\phi, a_1, a_2) = (0, 0, 0)$. There are two possible cases to consider: (1) $k_1^n < k_2^n$ and (2) $k_1^n \geq k_2^n$.

Assume to the contrary that $k_1^n < k_2^n$. In this case, the first k_1^n entrants (and only them) buy a license. Moreover, agent $i_{k_1^n}$ is indifferent whether to buy a license as, otherwise, the SO could increase his profit by lowering b_1. Furthermore, $b_2 > 0$ as, otherwise, distributors would never recruit. By inequalities (15) and (16), for a sufficiently large n, $w(b_1)/\kappa(b_1) > w(b_2)/\kappa(b_2)$. Hence, we can reduce the SO's expected cost given that the first k_1^n entrants purchase a license, by lowering b_2 by a small $\epsilon > 0$ and raising b_1 by $\epsilon w(b_2)/w(b_1)$. For small $\epsilon > 0$, this change has no effect on the cutoff k_1^n: up to that period agents find it optimal to purchase a license and from that period on they find it suboptimal to do so. Note that if $\epsilon > 0$ is small, then this change has no effect on the optimality of recruiting new distributors up to period k_1^n. Hence, the scheme that is generated by the changes in b_1 and b_2 induces an ABEE in which the SO's revenue is the same as in R^n (as the same agents purchase licenses) but his costs are lower. This is in contradiction to R^n being profit-maximizing.

So far, we have shown that, for large n, it must be that $k_1^n \geq k_2^n$ (as in Theorem 1). We now show that given a 2-level scheme R in which $a_1 = a_2 = 0$, any profile of strategies σ in which $k_1 \geq k_2$ is part of an ABEE of $\Gamma(R)$ if and only if it is part of a PBE of $\Gamma(R)$. Consider such a profile and the marginal agent i_{k_1}. In an ABEE (resp., PBE), he does not expect to make offers. He therefore correctly predicts an ABEE (resp., PBE) payoff of $v_{k_1}b_1 - c \geq 0$ if he purchases a license. Agents who enter in period $t > k_1$ correctly expect an ABEE (resp., PBE) payoff of $v_t b_1 - c < 0$ if they purchase a license, and agents who enter prior to k_1 expect to obtain more than $v_{k_1}b_1 - c \geq 0$ as they can refrain from selling licenses. Now consider a distributor who is supposed to recruit an agent i_t. Since $\mathbf{a} = \mathbf{0}$, in an ABEE (resp., PBE), the distributor correctly expects $b_2 v_t - \hat{c}$ for that recruitment regardless of whether i_t will recruit new members. We conclude that agents' behavior in σ is optimal given analogy-based expectations if and only if it is optimal given correct expectations.

We can combine the equivalence between ABEE and PBE given $k_1^n \geq k_2^n$ and 2-level schemes in which $(\phi, a_1, a_2) = (0, 0, 0)$ together with Theorem 1 to establish that, for a sufficiently large n, a scheme is profit-maximizing when agents are analogy-based reasoners if and only if it is profit-maximizing when agents are fully rational. □

Proof of Proposition 3. By Theorem 1, when the SO faces fully rational agents, there is a profit-maximizing scheme R that is a 2-level scheme with $(\phi, a_1, a_2) = (0, 0, 0)$ and, in its induced game, $k_1 \geq k_2$. In Step 5 of Proposition 2, we showed that, given a 2-level scheme in which $(\phi, a_1, a_2) = (0, 0, 0)$, a profile of strategies in which $k_1 \geq k_2$ is part of a PBE if and only if it is part of an ABEE. Hence, $\pi^{\text{ABEE}}(R) = \pi^{\text{PBE}}(R)$. □

References

Abreu, Dilip and Markus K. Brunnermeier (2003), "Bubbles and crashes." *Econometrica*, 71, 173–204. [636, 637]

Antler, Yair (2018), "Multilevel marketing: Pyramid-shaped schemes or exploitative scams." Working Paper 13054, CEPR. [645]

Antler, Yair (2021), "Multilevel marketing: Pyramid-shaped schemes or exploitative scams." Working Paper, Retrieved from yairantler.com. [654, 655]

Babaioff, Moshe, Shahar Dobzinski, Sigal Oren, and Aviv Zohar (2012), "On Bitcoin and red balloons." In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 56–73. [637, 639]

Bhattacharya, Prasad and Joseph Gastwirth (1984), "Two probability models of pyramid or chain letter schemes demonstrating the unreliability of the claims of promoters." *Operations Research*, 32, 527–536. [639]

Bianchi, Milo and Philippe Jehiel (2010), "Bubbles and crashes with partially sophisticated investors." Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1785002. [636]

Bort, Ryan (2016), "John Oliver says multilevel marketing companies like Herbalife are pyramid schemes." *Newsweek*, November 7. [633]

Brunnermeier, Markus K. and Jonathan A. Parker (2005), "Optimal expectations." *American Economic Review*, 95, 1092–1118. [654]

Direct Selling Association (2014), "U.S. direct selling data." Retrieved from https://www.dsa.org/docs/default-source/research/researchfactsheet2007-2014.pdf?sfvrsn=9e7ecea5_0. [634]

Drmota, Michael (2009), *Random Trees*. Springer, Vienna. [639]

Emek, Yuval, Ron Karidi, Moshe Tennenholtz, and Aviv Zohar (2011), "Mechanisms for multi-level marketing." In *Proceedings of the 12th ACM Conference on Electronic Commerce*, 209–218. [637, 639]

Eyster, Erik and Matthew Rabin (2005), "Cursed equilibrium." *Econometrica*, 73, 1623–1672. [636, 654]

Federal Trade Commission (2016), "Herbalife will restructure its multi-level marketing operations and pay 200 million for consumer redress to settle FTC charges." Retrieved from https://www.ftc.gov/news-events/news/press-releases/2016/07/herbalife-will-restructure-its-multi-level-marketing-operations-pay-200-million-consumer-redress. [633]

FTC v. Fortune Hi-Tech Marketing, Inc. 13 CV 578 (N.D. Ill, 2013). [634]

Gastwirth, Joseph (1977), "A probability model of a pyramid scheme." *American Statistician*, 31, 79–82. [639]

Heidhues, Paul, Botond Kőszegi, and Takeshi Murooka (2017), "Inferior products and profitable deception." *Review of Economic Studies*, 84, 323–356. [652]

Heidhues, Paul and Botond Kőszegi (2018), "Behavioral industrial organization." In *Handbook of Behavioral Economics*, Vol. 1, Chapter 6, 517–612. [636]

Jehiel, Philippe (2005), "Analogy-based expectation equilibrium." *Journal of Economic Theory*, 123, 81–104. [635, 636, 644, 647]

Keep, William and Peter Vander Nat (2014), "Multilevel marketing and pyramid schemes in the United States. An historical analysis." *Journal of Historical Research*, 6, 188–210. [634]

Koehn, Daryl (2001), "Ethical issues connected with multilevel marketing." *Journal of Business Ethics*, 29, 153–160. [637]

Mahmoud, Hosam (1991), "Limiting distributions for path lengths." *Probability in the Engineering and Informational Sciences*, 5, 53–59. [663]

Moinas, Sophie and Sebastien Pouget (2013), "The bubble game: An experimental study of speculation." *Econometrica*, 81, 1507–1539. [635, 636, 637, 647]

Piccione, Michele and Ariel Rubinstein (2003), "Modeling the economic interaction of agents with diverse abilities to recognize equilibrium patterns." *Journal of the European Economic Association*, 1, 212–223. [636]

Rantala, Ville (2019), "How do investment ideas spread through social interaction? Evidence from a Ponzi scheme." *Journal of Finance*, 74, 2349–2389. [636]

SEC (2013), "Beware of pyramid schemes posing as multi-level marketing programs." Retrieved from https://www.sec.gov/oiea/investor-alerts-bulletins/investor-alerts-ia_pyramidhtm.html. [634, 653]

Shiller, Robert J. (2015), *Irrational Exuberance*. Princeton University Press, Princeton. [636]

Spiegler, Ran (2011), *Bounded Rationality and Industrial Organization*. Oxford University Press, New York. [636]

Spiegler, Ran (2016), "Bayesian networks and boundedly rational expectations." *Quarterly Journal of Economics*, 131, 1243–1290. [654]

Stivers, Andrew, Douglas Smith, and Ginger Z. Jin (2019), "The alchemy of a pyramid: Transmutating business opportunity into a negative sum wealth transfer." SSRN Working Paper. [636]

Tirole, Jean (1982), "On the possibility of speculation under rational expectations." *Econometrica*, 50, 1163–1182. [637, 641]

Vander Nat, Peter and William Keep (2002), "Marketing fraud: An approach for differentiating multilevel marketing from pyramid schemes." *Journal of Public Policy & Marketing*, 21, 139–151. [637]

World Federation of Direct Selling Associations (2018), "Annual report 2017/2018." Retrieved from https://www.wfdsa.org. [633]

Co-editor Simon Board handled this manuscript.

Manuscript received 27 May, 2021; final version accepted 5 July, 2022; available online 15 July, 2022.