Multilevel Marketing: Pyramid-Shaped Schemes or Exploitative Scams?*

Yair Antler[†]

Abstract

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 a reward scheme that charges an entry fee and compensates each participant based on the number of people that he recruits and that these recruits recruit. By contrast, when the demand for a company's product is high, optimal multilevel marketing schemes neither charge entry fees nor pay directly for recruitment.

What delineates pyramid scams from legitimate multilevel marketing enterprises? Recent growth[1] in the multilevel marketing (MLM) industry—which over the past five years has engaged over 20 million[2] Americans—has raised the urgency of this question for consumer protection agencies. 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 take advantage of vulnerable individuals.

*I am grateful to Rani Spiegler for many insightful discussions and suggestions. I thank Vellore Arthi, Benjamin Bachi, Kfir Eliaz, Erik Eyster, Michael D. Grubb, Paul Heidhues, Philippe Jehiel, Alessandro Lizzeri, Francesco Nava, Ronny Razin, Santiago Oliveros, Michael Richter, and Ariel Rubinstein for useful comments and suggestions. 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, Royal Holloway, Technion, Tel Aviv, Toulouse, UCL, Warwick, and Zurich. I acknowledge financial support from the Henry Crown Institute of Business Research in Israel.

†Tel Aviv University, Ramat Aviv, Tel Aviv, 69978, Israel. *E-mail:* yair.an@gmail.com

[1]Membership in MLMs 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).

[2]According to the Direct Selling Association's (DSA) annual report (DSA, 2016).

The MLM industry's questionable legitimacy received considerable media attention[3] following a recent FTC investigation against Herbalife (FTC, 2016a). Identifying whether a particular company is a legitimate one, or whether it is an exploitative pyramid scam that promotes useless products and services in order 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 the distributors are encouraged to sell the product, and it is an illegal pyramid scam if it prioritizes recruitment over selling (FTC, 2016b; 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 compensation schemes. Our premise is that the potential distributors are strategic, and that the MLM company chooses a compensation scheme while taking these prospective distributors' incentives into account. To understand the structure of potential reward schemes, consider the following example.

Example 1 *The reward scheme R pays every distributor a commission of b_1 for every sale that he makes and a commission of a_1 for every agent he recruits to the sales force. The reward scheme R' pays every distributor a commission of b'_1 for every sale he makes and a commission of b'_2 for every sale made by one of his recruits. It also pays every distributor a'_1 for every person he recruits and a'_2 for every one of his recruits' recruits. Both schemes charge a license fee[4] of $\phi \geq 0$ from every distributor. We refer to a_1, a'_1, and a'_2 as recruitment commissions, and to reward schemes such as R (respectively, R') as 1-level (respectively, multilevel) schemes as they compensate the distributors based on the first level (respectively, multiple levels) of their downline.*

Both R and R' compensate the distributors for recruiting others to work for the company. In practice, however, over 90% of the network marketing industry uses multilevel schemes such as R' (DSA, 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[5] pyramid scams used multilevel reward schemes. What

[3] See, e.g., McCrum (2016), McKown (2017), Multi-level Marketing in America (2015), Moyer (2018), Parloff (2016), Pierson (2017), Suddath (2018), Truswell (2018), and Wieczner (2017).

[4] In practice, fees are often presented as training costs or a requirement to purchase initial stock.

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

can explain these stylized facts? Can a "legitimate" company benefit from charging license fees, paying recruitment commissions, or offering multiple routes through which individuals can join the sales force? Does the answer depend on whether the company promotes genuine goods or just the opportunity to recruit others to the sales force?

To address the above questions, we develop a model in which a scheme organizer (SO) sells a good to a network of agents that is formed randomly and sequentially. The agents buy only from people to whom they are directly connected. The SO uses a reward scheme to incentivize agents to sell the good and recruit others (sell distribution licenses) in order to reach new pools of customers, that is, agents with whom he has no direct link. A key feature of the model is that each agent's likelihood of meeting new entrants (i.e., potential buyers and distributors) decreases as time progresses, 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 others 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.[6] Hence, to better understand such scams and their underlying reward schemes, we depart from the rational expectations paradigm.

In the main part of the paper, we use the analogy-based expectation equilibrium (ABEE) framework (Jehiel, 2005) to relax the requirement that agents have 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 under-appreciate 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 expected 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 recruitment. In other words, there exists no 1-level scheme such that the SO makes a strictly positive expected payoff

[6]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).

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—e.g., 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.

We further investigate the features of cognitive biases that help sustain fraudulent scams by incorporating into the framework several behavioral models in which agents' beliefs are not necessarily statistically correct. In all of these models, multiple levels of recruitment commissions can facilitate scams. For example, in Section 6, we adapt Brunnermeier and Parker's (2005) model of self-deception to the framework. We find that even when agents deceive themselves into holding extremely overoptimistic beliefs about their location in the game tree (e.g., that they are likely to be "at the top of the pyramid") and, as a result, it is possible for the SO to sustain a pyramid scam by means of a 1-level scheme, the scam that maximizes the SO's expected profit is built on a multilevel scheme. The reason for this effect is that the agents overestimate the number of downline recruits more than they overestimate the number of people they directly recruit in this behavioral model.

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 a license fee, nor does it pay recruitment commissions. 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 large, the optimal scheme looks just like when agents are fully rational. However, when the demand is sufficiently small, the optimal scheme charges a license fee and pays for at least two levels of recruitment. Thus, the tools that pure pyramid scams are based on—recruitment commissions and license fees—disappear when there is a large demand for the good and emerge again when the demand is small.

Our results connect the SO's ability to scam the agents to the use of reward schemes that charge license fees and pay recruitment commissions. We study the implications of banning these tools within the model and show that such a regulation can reduce the profit of an SO who faces analogy-based reasoners. However, 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 when facing fully rational agents.

Related literature

Our paper relates to a strand of the behavioral industrial organization literature that studies the market settings and contractual features that enable firms to exploit agents who are subject to different biases.[7] Spiegler (2011) offers a textbook treatment of such models. In Eliaz and Spiegler (2006, 2008), a principal interacts with agents who differ in their ability to predict their future tastes. In Gabaix and Laibson (2006), firms may hide information about add-on prices from unaware consumers. Grubb (2009) studies contracting when agents are overconfident about the accuracy of their forecasts of their own future demand. Crawford et al. (2009) and Jehiel (2011) study manipulative auction design. Heidhues and Kőszegi (2010) study exploitative credit contracts when consumers are time-inconsistent. Grubb (2015) illustrates how various contractual features can be used to exploit overconfidence (e.g., automatic renewal). While this literature is not insubstantial, the various contractual features inherent in MLM are not fully understood.

We use the analogy-based expectation equilibrium (Jehiel, 2005) as our 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, agents fail to realize the extent to which the other players' actions depend on their private information. Piccione and Rubinstein (2003) study intertemporal pricing when consumers reason in terms of a coarse representation of the correct equilibrium price distribution. Other prominent models in which players' reasoning is coarse are Jehiel

[7] A related strand of the literature studies the vulnerability of nonstandard preferences or of non-standard decision-making procedures to exploitative transactions ("Dutch Books"). For example, see Laibson and Yariv (2007) and Rubinstein and Spiegler (2008).

and Koessler (2008), Mullainathan et al. (2008), Eyster and Piccione (2013), Guarino and Jehiel (2013), and Steiner and Stewart (2015).

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.[8] 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.[9]

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 enterprises have received considerable attention outside of the economics literature. A strand of the computer science literature (see, e.g., Emek et al., 2011; Babaioff et al., 2012) focuses on MLM mechanisms' 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; Keep and Vander Nat, 2002).

The paper proceeds as follows. The model is presented in Section 1. Sections 2 and 3 provide benchmark results about the social optimum and the SO's optimal scheme when agents are fully rational. Section 4 investigates pure pyramid scams and Section

[8]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 in order 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).

[9]For other prominent behavioral theories of bubbles see Harrison and Kreps, (1978) and DeLong et al. (1990).

5 studies MLM using ABEE. Section 6 examines the implications of leading behavioral models for our results, and Section 7 concludes.

1 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 $I = \{1, ..., n\}$. Each agent $i \in I$ has a unit demand and a willingness to pay $\omega_i \in \{0, 1\}$ that is drawn by nature when i enters the game. For each $i \in I$, denote $q := Pr(\omega_i = 1)$.

In each period $t = 1, 2, ..., n$, one agent enters the game. We refer to the t-th entrant as agent i_t. Upon entering the game, agent i_t meets one player $j \in \{SO, i_1, ..., i_{t-1}\}$ chosen uniformly at random by nature.[10] 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 $D_1 = \{SO, i_1\}$ if i_1 accepts an offer and $D_1 = \{SO\}$ 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. 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 $i \in I \cup \{SO\}$ to $j \in I$ if such a path exists.

To illustrate the game tree, suppose that, at the end of period 9, G is as presented in Figure 1 and $D_9 = \{SO, i_1, i_2, i_3\}$. Recall that agent i_{10} is equally likely to meet each $j \in \{SO, i_1, ..., i_9\}$. 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.

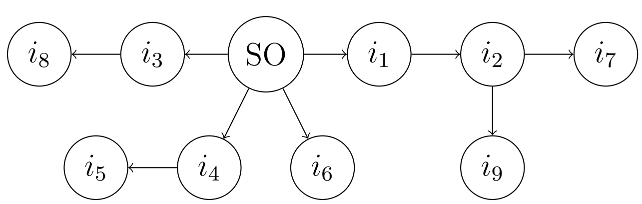

Figure 1: A snapshot of G at the end of period 9.

Agents in the model make two types of strategic choices: whether to accept an offer

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

upon entering the game, and, conditional on purchasing a license, whether to make offers when they meet agents later in the game. The choice of whether to purchase the good for personal consumption is modeled in a nonstrategic manner, and will be clarified in the next paragraph. In each period t, every player $i \in \{SO, i_1, ..., i_t\}$ 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 accepted an offer and whether j bought the good. Denote by H_i the set of information sets in which player $i \in I \cup \{SO\}$ 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 $\sigma = (\sigma_i)_{i \in I \cup \{SO\}}$ to denote a profile of strategies.

Distributors are paid according to a reward scheme that consists of:

- A license fee $\phi \geq 0$.
- Recruitment commissions: $a_1, a_2, a_3, ... \geq 0$.
- Sales commissions: $b_1, b_2, b_3, ... \geq 0$.
- A price $\eta > 0$ for which each unit of the good is sold.

When agent i_t accepts an offer, he pays ϕ to the SO and the SO pays a commission of $a_{d(l,i_t)}$ to every distributor l on the directed path from the SO to i_t. We assume that, for each $t \in \{1, ..., n\}$, upon entering the game, agent i_t purchases a unit of the good for personal consumption if and only if both (i) $\omega_{i_t} \geq \eta$ and (ii) i_t's immediate predecessor in G is a member of D_{t-1} (i.e., a distributor or the SO). If agent i_t purchases the good, then he obtains an additional utility of ω_{i_t}, pays η to the SO, and the SO pays a commission of $b_{d(l,i_t)}$ to every distributor l on the directed path from the SO to i_t. In addition to the commissions, when an agent accepts an offer, he incurs a cost of $c > 0$, which reflects the cost of learning how to market the good, and the person who recruited him incurs a cost of $\hat{c} > 0$, which reflects the cost of training the new distributor.[11]

We denote by $\Gamma(R)$ the game that is induced by the reward scheme R. Agents 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 $\pi(R)$.

The SO faces the risk that the distributors will create fictitious recruits in order

[11]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.

to become eligible for additional commissions.[12] Motivated by this risk, 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 incentive compatibility constraint rests on the assumption that the SO can verify[13] 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.

Discussion: 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).[14] This process rests on the assumptions that there is a deterministic date at which the game ends and that each agent is directly connected to only one agent upstream. Our main results do not depend on these assumptions. Nonetheless, we use this process since it allows us to convey the main messages while keeping the exposition simple. As we show in the Supplemental Appendix, the main insights hold when there is uncertainty about the length of the game and when agents can be connected to multiple agents upstream.

2 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 only if they meet a distributor upon entering the game. In expectation, agent i_t meets

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

[12]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).

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

[14]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. In none of these papers, however, is there strategic interaction.

new entrants in periods $t+1,...,n$ (for completeness, let $v_n := 0$). Thus, if agent i_t becomes a distributor he can provide the good to v_t agents who would not consume it otherwise, where q of them are willing to pay 1 for the good. Hence, the direct social benefit from turning agent i_t into a distributor is qv_t. This cost-benefit analysis leads to the following definition.

Definition 1 *A reward scheme R is said to be socially optimal if there exists an equilibrium of $\Gamma(R)$ in which, in every period $t \in \{1,...,n\}$, if $qv_t > c+\hat{c}$, then the entrant purchases a license and, if $qv_t < c+\hat{c}$, then he does not purchase a license.*

Note that Definition 1 takes into account only direct sales. 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 sales. 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. Thus, when looking for a socially optimal scheme we can ignore downline sales and recruitment.

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. A scheme R is said to be profit-maximizing if there exists no IC scheme R' such that $\pi^{PBE}(R') > \pi^{PBE}(R)$. 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 indifferent whether to make an offer or not makes it.[15] To avoid trivial cases in which the reward scheme is essentially irrelevant, we shall assume in this section that n is sufficiently large such that in every PBE of a profit-maximizing scheme at least two agents purchase licenses. In the Supplemental Appendix, we show that this requirement is satisfied when $qv_4 > c+\hat{c}$, where v_4 is given in (1) for $t=4$.

[15] 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.

Consider 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.[16]

Lemma 1 *Consider a reward scheme R in which $\eta \leq 1$. In a PBE of $\Gamma(R)$:*

1. *If agent i_t receives an offer, he accepts it if and only if $v_t q b_1 \geq c + \phi$.*

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

Lemma 1 establishes that agents' PBE behavior depends only on ϕ, b_1, a_1, and b_2. Hence, 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 schemes in which $a_2 = a_3 = ... = 0$ and $b_3 = b_4 = ... = 0$.

Having considered the agents' perspective, we now turn to study the scheme that maximizes the SO's expected payoff.

Theorem 1 *If R is IC and profit-maximizing, then $\phi = 0 = a_1$.*

To understand why charging a license fee is detrimental to the SO, consider the last agent who is supposed to buy a license in a PBE and denote him by k. At the optimum, agent k is indifferent whether to buy a license or not as, otherwise, the SO could increase his profit by charging a higher license fee without affecting the agents' equilibrium behavior. Suppose that the SO eliminates the license fee and, at the same time, lowers the commission b_1 such that agent k remains indifferent. This exercise has no effect on the expected net transfers from the SO to agent k. However, it reduces the expected net transfers the SO makes to agents who buy a license prior to agent k. The reason is that the reduction in b_1 has a stronger effect on agents who make more sales while the reduction in ϕ is independent of sales. In other words, scaling down ϕ and

[16] The lemma assumes that $\eta \leq 1$ as, otherwise, agents do not purchase the good and $\pi^{PBE}(R) = 0$.

b_1 reduces the rents agents obtain by purchasing a license early in the game. Thus, by eliminating the fee and reducing b_1 the SO increases his expected payoff.

Observe 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 commissions. As we show in the proof, this effect is of second order such that the SO is better off using a scheme that does not charge a license fee and does not pay for recruitment.

We can conclude that to maximize his expected payoff, the SO pays b_1 for direct sales and b_2 for indirect sales. As Lemma 1 shows, the agents' behavior 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, distributors recruit every agent they meet up to period k_2, and the SO is the only one making offers in periods $k_2 + 1, ..., k_1$. In a profit-maximizing scheme, agent i_{k_1} must be indifferent whether to accept an offer or not and distributors must be indifferent whether to recruit agent i_{k_2} or not. Hence, by Lemma 1, $b_1 = \frac{c}{qv_{k_1}}$ and $b_2 = \frac{\hat{c}}{qv_{k_2}}$ in a profit-maximizing scheme.

Lemma 1 also allows us to find socially optimal schemes. By Definition 1, if $m = max\{t|qv_t \geq c + \hat{c}\}$, then every scheme where the first m entrants (and only them) purchase a license in a PBE of its induced game is socially optimal. By Lemma 1, any scheme that pays $b_1 = \frac{c+\phi}{qv_m}$, $b_2 = \frac{\hat{c}-a_1}{qv_m}$, and that charges $\eta = 1$, incentivizes this behavior. In particular, there exists a socially optimal scheme $\hat{R}$ that pays $\hat{b}_1 = \frac{c}{qv_m}, \hat{b}_2 = \frac{\hat{c}}{qv_m}$, $\hat{a}_1 = 0$, and that charges $\hat{\eta} = 1$ and $\hat{\phi} = 0$. Since $qv_m \geq c + \hat{c}$, $\hat{R}$ is IC. The next corollary summarizes this finding.

Corollary 1 *There exists an IC socially optimal scheme that does not charge a license fee and does not pay recruitment commissions.*

In the Supplemental Appendix (Proposition 13), we show that profit-maximizing schemes are not socially optimal unless n is small. In general, profit-maximizing schemes pay lower commissions and incentivize fewer agents to become distributors than is socially optimal. From the SO's perspective, increasing the number of distributors from $t - 1$ to t entails paying higher commissions that make it worthwhile for i_t to buy a license. Since all agents are paid according to the same reward scheme, by increasing the commissions 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.

The results in this section show that the SO does not benefit from using recruitment commissions and charging license fees and that he need not use these tools to maximize social welfare when facing fully rational agents. This might suggest that if license fees and commissions are used, then some kind of agent irrationality is involved. In the following sections, we study settings in which agents are boundedly rational and investigate in which situations these components of the reward scheme can be useful both from the SO's perspective and from the social perspective.

4 Pure Pyramid Scams ($q = 0$)

To capture the idea that the main "product" 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 products that are being traded in the model are distribution licenses.[17] Intuitively, such a market should not exist as trade in distribution licenses does not add any value. Indeed, at the social optimum, no agent purchases a license in this case.

If $q = 0$ and there exists a scheme R such that the SO makes a strictly positive expected profit in an equilibrium of its induced game, then we say that the SO can *sustain a pure pyramid scam*. Proposition 1 establishes that when all of the agents are fully rational, he cannot do so.

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

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).

Our main objective is to understand the forces and compensation plans that enable pyramid scams to operate. As Proposition 1 shows, it is impossible to do so by means of the classic rational expectations model and we shall therefore depart from this model.

4.1 The baseline behavioral model

Jehiel (2005) suggests an elegant framework that incorporates partial sophistication into extensive-form games. In this framework, different contingencies are bundled into

[17] As will become clear later, the main insights we obtain in this section remain valid when $q > 0$ as long as q is small with respect to c and $\hat{c}$.

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.

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 $M_1 := \cup_{i \in I} H_i^1$ and $M_2 := \cup_{i \in I} H_i^2$. 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 $i \in I$, $\beta^i = (\beta_1^i, \beta_2^i)$ are agent i's analogy-based expectations about the other agents' behavior. A strategy σ_i is a best response to β^i if it is optimal given a belief that each agent $j \neq i$ accepts every offer he receives with probability β_1^i and that, if j has the opportunity to make an offer, then he makes it with probability β_2^i. Let $\beta := (\beta^i)_{i \in I}$.

Definition 2 *Agent i's analogy-based expectations β^i are said to be consistent with the profile of strategies σ if, for every $k \in \{1, 2\}$, it holds that $\beta_k^i = \frac{\sum_{h \in M_k} r_\sigma(h)\sigma(h)}{\sum_{h \in M_k} r_\sigma(h)}$ whenever $r_\sigma(h) > 0$ for some $h \in M_k$.*

Definition 3 *The pair (σ, β) forms an analogy-based expectation equilibrium (ABEE) if, for each $i \in I$, β^i is consistent with σ and σ_i is a best response to β^i.*

Consistency implies that, in an ABEE, $\beta_1^i = \beta_1^j$ and $\beta_2^i = \beta_2^j$ for every pair of agents $i, j \in I$. Therefore, we shall omit the superscript and use β_1 and β_2 instead.

Discussion: Consistency, analogy classes, and the SO's strategy

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 β_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 $i \in D_{t-1}$ 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, $\beta_1 = \frac{1}{1+1+\frac{2}{3}} = \frac{3}{8} > \frac{1}{3}$ is consistent with σ.

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. 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 players (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.

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 commitment as his behavior affects β. The SO's commitment power allows us to simplify the exposition, but does not affect the results.

4.2 Structure and existence of pyramid scams

The main result of this section (Theorem 2) shows that the SO cannot sustain a pyramid scam by means of a scheme that pays the distributors only for the number of licenses they sell. Thus, behind every pyramid scam is a scheme that pays for at least two levels of recruitment. Theorem 3 shows that if the number of potential participants is sufficiently large, a pyramid scam is sustainable and the necessary condition of Theorem 2 is tight. Finally, Theorem 4 shows that the SO may have to use more than two levels of compensation to sustain a pyramid scam.

We start by showing that the SO cannot sustain a pyramid scam by means of a reward scheme that compensates distributors only for the number of agents they directly recruit. We shall refer to such schemes as 1-level schemes.

Definition 4 *A reward scheme R is said to be a z-level scheme if $a_\tau = 0$ and $b_\tau = 0$ for every $\tau > z$.*

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

To gain intuition for this result, let us suppose for a moment that agents purchase licenses in an ABEE and study their analogy-based expectations. Since the likelihood of meeting new entrants decreases over time, there is a period t such that agents accept offers up to period t and reject offers afterward. Nonetheless, distributors continue making offers after period t because they falsely believe that the other agents might accept them. In fact, each distributor tries to recruit every person he meets,[18] and so makes, in expectation, v_t offers after period t, where v_t is as given in (1). Thus, every offer that is accepted in periods $1, ..., t$ results in a distributor who makes, in expectation, v_t offers after period t, which are all rejected. We can conclude that the proportion of accepted to total offers, β_1, cannot exceed $\frac{1}{1+v_t}$.

Imagine agent i_t, the last agent who is supposed to purchase a license in our ABEE, contemplating an offer. Conditional on accepting, he expects to make v_t offers and falsely believes that each of them will be accepted with probability β_1. The more offers he expects to make, the more offers are made (and rejected) late in the game, which implies that β_1 is lower. Overall, i_t falsely expects to recruit $v_t\beta_1$ agents, which is less than 1 regardless of how large v_t is. If the scheme is IC, then $a_1 \leq \phi$, which means that no commission on his own sales would cover the fee paid. Hence, an ABEE in which agents purchase licenses cannot exist in a 1-level scheme's induced game.

The next result establishes that the SO can sustain a pyramid scam by means of a 2-level scheme if n is large.

Theorem 3 *Let $q = 0$. There exists an integer $n^\star$ such that an IC 2-level scheme R for which $\pi^{ABEE}(R) > 0$ exists if and only if $n \geq n^\star$.*

[18] If $a_1 \geq \hat{c}$, distributors always find it optimal to recruit the new entrant. Note that, in a 1-level scheme, if $a_1 < \hat{c}$, then purchasing a license leads to a negative profit no matter how many individuals one recruits, and so our conjectured ABEE cannot exist.

The intuition for the existence result resembles the one behind 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 β_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 recruits their recruits will recruit in the future. The agents' imperfect perception of the other agents' behavior leads them to overestimate both of these variables. As we showed in Theorem 2, agents whose beliefs are statistically correct do not overestimate their 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, the agents do not overestimate their recruits' ability to sell licenses by much. However, the accumulation of agents' mistakes allows the SO to overcome the incentive-compatibility constraint and sustain a pyramid scam.

In light of Theorem 3 it is elementary to ask whether there are cases in which the SO must use more than two levels of compensation. The next result illustrates that when n is small, the answer is affirmative.

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

Generality of the results: A road map

Throughout this section, we imposed three main assumptions. First, we assumed that $q = 0$, which made it salient that trade is inefficient. Second, we used a simple network formation process, which allowed us to abstract from situations in which agents are connected to multiple agents upstream and in which there are cycles in the social network. Finally, we assumed that agents' beliefs are statistically correct, which constrained the overoptimism of the agents and the SO's ability to exploit them. We now relax these assumptions. In Section 5, we let $q > 0$ and show that the profit-maximizing schemes resemble the ones identified in this section when the demand is low, but when the demand is high, the profit-maximizing schemes resemble the ones we identified in the rational expectations model of Section 3. In the Supplemental Appendix, we examine different models of social networks and show that the main results of this section hold. Finally, in Section 6, we incorporate leading behavioral models into our framework and provide additional support to the finding that multiple levels of recruitment commissions can be a litmus test for exploitative MLM contracts.

5 Multilevel Marketing of Genuine Goods ($q > 0$)

We now examine an environment where the good is intrinsically valued and the agents are vulnerable to deceptive practices. Thus, the SO can benefit both from the agents' sales and from their mistakes. We start by studying a setting in which pyramid scams are viable and the demand for the good is low (fixed $n > n^\star$, small $q > 0$). We show that profit-maximizing schemes charge license fees and pay for multiple levels of downline recruits, as the pure pyramid scams in Section 4. We then turn to the case where the demand is large (fixed $q > 0$, large n) and show that, as in the rational expectations case of Section 3, profit-maximizing schemes do not charge a license fee and do not pay for recruitment.

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 can benefit from raising the stakes of the contract. To guarantee that a profit-maximizing scheme exists, we shall assume that the maximal amount that each agent can pay for a license is B and that B is large w.r.t. c and $\hat{c}$. Second, to simplify the analysis, we fix an arbitrary integer $\tau^\star > 2$ and restrict attention to $\tau^\star$-level schemes; i.e., we impose that $a_\tau = b_\tau = 0$ for every $\tau > \tau^\star$.

We start the analysis by showing that when q is sufficiently small, then profit-maximizing schemes charge a license fee and pay for recruitment.

Proposition 2 *For every $n > n^\star$, there exists a number $q^\star(n) > 0$ such that if $q < q^\star(n)$, then every IC profit-maximizing scheme charges a license fee and pays for at least two levels of downline recruits.*

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 or charge license fees are not profit-maximizing.

Proposition 3 *Fix $q > 0$. There exists a number $\hat{n}$ such that if $n > \hat{n}$, then, in every IC profit-maximizing scheme, $a_1 = a_2 = ... = b_3 = b_4 = ... = \phi = 0$, $b_1 > 0$, and $b_2 > 0$.*

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 profit from sales: a fee makes it more costly to become a distributor, which requires paying higher commissions to attract prospective distributors. This effect becomes stronger when the potential profit from sales qn is large. In such instances, the SO uses many distributors to increase the number of agents who purchase the good. Thus, he has to pay multiple commissions for every sale and every recruitment such that raising the commissions to compensate for the license fee becomes extremely costly and the SO is better off not charging a license fee. Note that an IC scheme that does not charge a license fee cannot pay for recruitment either. Thus, profit-maximizing schemes do not charge license fees and do not pay recruitment commissions.

Proposition 3 shows that when the demand is large, the SO cannot benefit from extending the scheme beyond level two. Note that this is different from the case in which q is small, where the SO may find it beneficial to do so (Theorem 4 shows this for $q = 0$ and can be extended to the case of small $q > 0$).

The main findings of this section are summarized in the following corollary.

Corollary 2 *There exists a number $\tilde{n}$ such that for every $n > \tilde{n}$:*

- *If q is sufficiently large, then in every IC profit-maximizing scheme $a_1 = ... = a_{\tau^\star} = b_3 = ... = b_{\tau^\star} = \phi = 0$.*
- *If q is sufficiently small, then in every IC profit-maximizing scheme $\phi > 0$, $a_1 > 0$, and $a_2 > 0$. Moreover, it is possible that $a_\tau > 0$ for $\tau > 2$.*

Banning recruitment commissions and license fees

Proposition 2 shows that there are instances in which banning recruitment commissions and license fees strictly reduces the SO's potential profit. The next result establishes a bound on this effect by showing that, even under such restrictions, the SO's potential expected payoff is no lower than his potential expected payoff when he is not restricted and agents are fully rational. The latter payoff can be viewed as the fundamental value of the operation.

Proposition 4 *For every scheme R, there exists an IC 2-level scheme $\tilde{R}$ such that $\tilde{\phi} = \tilde{a}_1 = \tilde{a}_2 = 0$ and $\pi^{ABEE}(\tilde{R}) \geq \pi^{PBE}(R)$.*

In Section 3, it was established that when agents are fully rational there is an IC profit-maximizing 2-level scheme $\tilde{R}$ in which $\tilde{a}_1 = \tilde{a}_2 = \tilde{\phi} = 0$. Proposition 4 shows that the agents' behavior in a PBE of $\Gamma(\tilde{R})$ is identical to their behavior in an ABEE of $\Gamma(\tilde{R})$ such that the SO's expected payoff is the same under both solution concepts.

To understand why analogy-based reasoners behave as if they were fully rational, note that these agents correctly predict (i) how many units of the good they will sell in the future and (ii) how many units their recruits will sell *conditional on buying a license*. In the ABEE/PBE of $\Gamma(\tilde{R})$, these are the only variables agents have to estimate in order to assess whether to buy a license and whether to recruit new entrants.

The social optimum

The IC 2-level scheme $\hat{R}$ that was introduced in Section 3 incentivizes the same socially optimal behavior both under ABEE and under PBE. The intuition and proof for this result are similar to those of Proposition 4. For brevity, they are omitted.

Corollary 3 *There exists an IC socially optimal scheme that does not charge a license fee and does not pay for recruitment.*

In general, IC socially optimal schemes are not profit-maximizing.[19] When qn is large, they induce more distributors and pay higher commissions than profit-maximizing schemes as in the baseline model of Section 3 and for similar reasons. When n is large but q is small such that qn is small, IC socially optimal schemes are not profit-maximizing either, but for a different reason: they induce fewer distributors than profit-maximizing schemes (whose main objective is to benefit from the fees these distributors pay). Indeed, Proposition 2 shows that even when q is small such that MLM

[19]See Section B.2.2 in the Supplemental Appendix.

is not socially optimal, profit-maximizing schemes can incentivize some of the agents to purchase a license. Thus, under ABEE, compared to the social optimum, when the demand is small profit-maximizing schemes induce too many distributors but when it is large they induce too few distributors.

Non-distributor consumers. When profit-maximizing schemes induce more distributors than socially optimal ones, distributors incur losses. However, it also means that more agents consume the good relative to the social optimum. While in our model, due to the binary demand, these consumers are left with no surplus, under a more realistic demand structure these consumers would enjoy a higher expected utility than at the social optimum. Thus, when the demand for the good is small, profit-maximizing schemes can leave the SO and the non-distributor consumers with a higher surplus than they would obtain at the social optimum at the expense of some of the distributors.

6 Alternative Behavioral Explanations

The main behavioral model used in the paper captures the idea that agents do not fully grasp how difficult it becomes to recruit new members over time. We now consider alternative models of distorted beliefs that capture psychological phenomena that can be relevant in the context of MLM, adapt them to this paper's framework, and study their implications for our results. We start with a model of motivated reasoning in which individuals derive anticipatory utility from expecting good future outcomes and deceive themselves into holding overoptimistic beliefs. We then turn to a model of cognitive hierarchies in which individuals believe that they understand the setting better than others. Finally, in the Supplemental Appendix, we study a model of social networks in which agents correctly predict the probability with which their direct friends accept offers but neglect the correlation between this probability and the number of friends their friends have. The main message in all of these models is that multiple levels of recruitment commissions facilitate scams.

Motivated Reasoning

Brunnermeier and Parker (2005) develop a model of motivated reasoning in which individuals derive anticipatory utility from expecting that good things will happen in the future.[20] In their model, individuals deceive themselves into holding (potentially) false

[20]For a comprehensive review of the idea of motivated reasoning see Bénabou and Tirole (2016).

beliefs, while balancing between anticipatory utility gains and the fact that distorted beliefs can lead to suboptimal choices in the future. Brunnermeier and Parker show that this type of reasoning leads to overoptimism and has significant implications in situations where returns are positively skewed, such as MLM.[21]

We now adapt Brunnermeier and Parker's behavioral model to our framework. In the spirit of their prescription, we assume that "each agent's beliefs are set taking as given the reaction functions of other agents." We simplify the model to the bare minimum by setting $\hat{c} = 0$ such that every distributor who meets an agent finds it optimal to recruit him. This allows us to posit that distributors always make an offer when they meet an agent and focus only on the agents' decisions whether to buy a license or not. It should be stressed, however, that the result presented in this section is unaffected by this assumption. To guarantee that a profit-maximizing scheme exists, we assume as in Section 5 that agents cannot pay more than $B > 0$ for a license.

We modify the baseline model by assuming that the order in which agents enter the game is drawn uniformly at random at the beginning of the game and that agents do not know the time at which they enter. When an agent receives an offer he updates his beliefs based on the identity of the proposer. We denote by $PR_i(t|j)$ the objective probability of i being the t-th entrant conditional on receiving an offer from $j \in I \cup \{SO\}$. On the equilibrium path, $PR_i(t|j)$ is determined according to Bayes' law.[22]

Agent i's strategy $s_i : \{SO, 1, ..., n\} \to \{0, 1\} \times \Delta(\{1, ..., n\})$ specifies a belief $\mathbf{r_i} \in \Delta(\{1, .., n\})$ and a purchasing decision $\sigma_i \in \{0, 1\}$ as a function of the player who made an offer to i. We use $r_i(t|j)$ to denote i's subjective belief that he is the t-th entrant conditional on receiving an offer from j. For each $i \in I$, let $s_{-i} = (s_j)_{j \in I - \{i\}}$ and denote agent i's expected rewards given that the other agents play s_{-i} and i purchases a license in period t from j by $p_t(s_{-i}, j)$. Let $\psi \in (0, 1)$ be the weight that captures the magnitude of the agents' anticipatory utility. A profile of strategies forms an equilibrium if for each $i \in I$ and $j \in I \cup \{SO\}$ agent i's strategy maximizes

$$\sigma_i(j) \left[(1 - \psi) \sum_{t=1}^{n} PR_i(t|j) p_t(s_{-i}, j) + \psi \sum_{t=1}^{n} r_i(t|j) p_t(s_{-i}, j) - \phi - c \right], \tag{2}$$

subject to the requirements that (1) $r_i(t|j) = 0$ whenever $PR_i(t|j) = 0$ and (2) $\sigma_i(j)$ maximizes

[21]See, e.g., Herbalife (2019).

[22]We refrain from placing restrictions on the agents' beliefs off the equilibrium path as the equilibria we analyze have full support.

$$\sigma_i(j)\Bigg[\sum_{t=1}^{n} r_i(t|j)p_t(s_{-i},j)-\phi-c\Bigg]. \tag{3}$$

To simplify the analysis, we shall focus on *symmetric equilibria*, that is, equilibria in which $s_i(k)=s_j(k)$ for every $i,j\in I$ and $k\in I\cup\{SO\}$.

Proposition 5 establishes that the potential profit of a scam is strictly larger with multilevel schemes than it is with 1-level schemes.

Proposition 5 *Set $q=0$. There exists $\tilde{n}$ such that for every $n>\tilde{n}$, there exist $\psi^\star(n)>0$ and $\psi^{\star\star}(n)>\psi^\star(n)$ such that:*

- *If $\psi<\psi^\star(n)$, then there exists no IC scheme R such that $\pi^{BP}(R)>0$.*
- *If $\psi\in[\psi^\star(n),\psi^{\star\star}(n))$, then there exists no IC 1-level scheme R such that $\pi^{BP}(R)>0$, but there exists an IC multilevel scheme R such that $\pi^{BP}(R)>0$.*
- *If $\psi\geq\psi^{\star\star}(n)$, then there exists an IC 1-level scheme R such that $\pi^{BP}(R)>0$. However, every scheme that is both 1-level and IC is not profit-maximizing.*

Proposition 5 relies on the fact that, in equilibrium, agents expect to recruit fewer agents than the total number of agents they expect their recruits to recruit. To see why, consider an agent who buys a license. He maximizes anticipatory utility by deceiving himself into believing that he entered the game first (or second) and, hence, that he is likely to recruit many agents. Since the likelihood of meeting new entrants goes down over time, he expects to do most of the recruiting early in the game. Thus, he expects to recruit mostly early entrants who are also likely to recruit many new members themselves. Overall, our "first" entrant falsely expects that he will recruit v_1 agents and that, in turn, these agents will recruit $\sum_{t=2}^{n-1}\frac{v_t}{t}>v_1$ agents. When $\psi<1$ this effect is mitigated by the agent's objective beliefs but does not disappear.

The above argument implies that, all else equal, our agent is better off if the SO increases a_2 and reduces a_1 by the same amount. For the SO, such a change saves costs: he pays a_1 for every distributor j such that $d(SO,j)>1$ and pays a_2 only if $d(SO,j)>2$. Thus, a 2-level scheme can induce a higher expected payoff for the SO and a higher perceived payoff for the agents. Hence, the SO can sustain a scam by means of a 2-level scheme when it is impossible to do so by means of a 1-level scheme (intermediate ψ), and can increase his expected profit by means of a 2-level scheme when it is possible to sustain a pyramid scam by means of a 1-level scheme (high ψ).

So far, we have established that the scope to overestimate the value of a distributorship and the potential profit of a scam increase when the SO moves from a 1-level to a 2-level scheme. It is possible to apply the same argument to show that, for every z, if n is sufficiently large, then there exists an IC $(z+1)$-level scheme that induces a strictly higher expected profit for the SO than every IC z-level scheme. That is, the scope to exploit the agents strictly increases with the number of levels.

In broader terms, the distribution of rewards in the game is positively skewed: agents who purchase a license early in the game can make a large profit while the average agent incurs a smaller loss. As shown by Brunnermeier and Parker (2005), the agents find this type of skewed distribution of rewards attractive: they deceive themselves into believing that high rewards are very likely. Essentially, this is what makes the agents pay for a license when ψ is not too small. Roughly speaking, increasing the number of levels of compensation enables the SO to increase the skewness of the rewards distribution, thereby increasing the scope to overestimate the value of a license and, as a result, increasing the potential profit of the scam.

Cognitive hierarchies

Level-k reasoning and cognitive hierarchy models are non-equilibrium models developed to capture the way individuals behave in novel strategic situations.[23] These models posit that players anchor their beliefs in a naive representation of the world that is supposed to capture how people would play the game instinctively. Individuals then adjust their behavior iteratively, performing a finite number of steps. We now adapt this approach to the context of MLM and pyramid scams, in which it makes sense to think that for many participants the situation is novel.

To focus on the main insights of Section 4, let $q = 0$. The baseline setting is modified by assuming that the cost $\hat{c}$ is paid in advance regardless of whether a distributor succeeds in recruiting an agent or not. This modification changes the interpretation of this cost, but does not change the benchmark results of Section 3, in which, in equilibrium, agents know who will accept their offers and who will reject them.

There are $k+1$ cognitive types $l_0, l_1, l_2, ..., l_k$. As in any model of level-k reasoning, the analysis inevitably hinges on the agents' naive model of the world, l_0. We assume that l_0 agents buy a license with probability $\alpha_0 > 0$ and, conditional on buying a license, make an offer to every agent they meet with probability $\alpha_1 > 0$. We assume that α_0

[23]See Crawford et al. (2013) for a comprehensive review.

is small with respect to B and $\hat{c}$ and, in particular, $\alpha_0 \leq \frac{\hat{c}}{B}$. As in models of level-$k$ reasoning, we assume that agents of cognitive type l_k best respond to a belief that all other agents are of type l_{k-1}. In the cognitive hierarchy approach of Camerer et al. (2004), agents of type l_k believe that other agents' types are distributed on $l_0, l_1, ..., l_{k-1}$. However, this distinction does not play a role in our setting as the key interaction is between types l_0 and l_1, whose behavior is identical under both approaches.

Type l_0's behavior is unaffected by the reward scheme. We now examine the behavior of higher cognitive types. We follow Crawford and Iriberri (2007) by assuming that "the anchoring L0 type exists mainly in the minds of higher types" that is, we assume that the support of the type distribution is $\{1, ..., k\}$.

Proposition 6 *There exists no IC 1-level scheme R such that $\pi^{LK}(R) > 0$. However, there exists a number $\overline{n}$ such that if $n > \overline{n}$, then there exists an IC 2-level scheme R such that $\pi^{LK}(R) > 0$.*

Proposition 6 shows that under our assumption on l_0, the threshold where pyramid scams start being viable is two levels of compensation. To see why, consider an agent of type l_1. In a 1-level scheme, he earns a_1 for every person he recruits and pays $\hat{c}$ for every offer he makes. Incentive compatibility implies that a_1 is bounded by $\phi \leq B$. Since our agent believes that other agents are of type l_0 and, therefore, that they accept offers with a probability of only $\alpha_0 \leq \frac{\hat{c}}{B}$, he thinks that making an offer yields a negative net expected return. Taking the fee into account, he prefers not to purchase a license.

In a 2-level scheme, an agent of type l_1 expects a_1 for every person he recruits and a *passive income* of a_2 for every individual that person recruits. Thus, our agent believes that recruiting an agent i_t yields a passive income of $\alpha_0 \alpha_1 v_t a_2$. When n is large w.r.t. t, our agent believes that i_t will sell many licenses and, therefore, that the passive income component will be large (i.e., v_t goes to infinity when n goes to infinity). Hence, for sufficiently large values of n, l_1 agents find it beneficial to purchase a license and to try and recruit new members up to some period $\tau_1(n)$. The larger n is, the larger $\tau_1(n)$ is.

Agents' behavior is monotone in their type. If l_1 agents reject every offer (as in IC 1-level schemes), then l_2 agents do not expect to recruit anyone and, therefore, reject every offer as well. This logic can be iterated for higher types. On the other hand, if l_1 agents purchase licenses and make offers to every agent they meet up to period $\tau_1(n)$ and $\tau_1(n)$ is sufficiently large, then agents of type l_2 find it optimal to purchase licenses and make offers up to some period $\tau_2(n) < \tau_1(n)$ as they perceive other agents to be of type l_1. This logic can be iterated for higher types as well. Hence, for large values of n, there is a sequence of cutoffs $\tau_1(n) > \tau_2(n) > ... > \tau_k(n) > 0$ such that l_j agents accept every offer up to period $\tau_j(n)$ and reject every offer made afterward.

7 Concluding Remarks

Legitimate MLMs and fraudulent pyramid scams are two widespread phenomena. Experts and potential participants often find it hard to distinguish between them. We developed a model that enables drawing a boundary between the two based on observable properties of their underlying reward schemes. 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. We illustrated that when agents hold distorted overoptimistic beliefs that are not statistically correct, it might be possible to sustain such a pyramid scam by means of a 1-level scheme; however, in these instances, maximizing the company's profit from a scam requires using a multilevel scheme. In all of these pyramid scams, the company lures the distributors by paying recruitment commissions and makes its profit from the fees they pay. The paper's benchmark results show that companies with a "good" product that face rational agents find it detrimental to use these two tools.

In the model, MLM enables companies that produce good products to incentivize agents to sell the product and recruit others to the sales force, thereby reaching pools of customers who would not purchase its product otherwise. It is natural to ask what is the advantage of MLM for such companies compared to more traditional marketing methods. MLM could be appealing in situations where there is a great deal of uncertainty about quality or fit. Consumers who are skeptical about a product's quality or whether it fits their specific needs would be reluctant to buy it at a high price. Traditional marketing channels may not allow the company to credibly communicate the merits of its product. However, when a consumer buys from a friend as in the MLM model, the latter might be better informed about the consumer's specific needs and the social capital in the interaction may allow him to credibly explain the merits of the product. Hence, MLM can outperform traditional marketing channels by decreasing uncertainty, thereby increasing the consumer's willingness to pay for the product.

It is possible to embed the baseline model in a simple framework in which the SO first learns q and then chooses whether to use MLM as in the baseline model or a more traditional marketing channel. The parameter q can be interpreted as the share of the population for whom the product fits. Ex ante, each agent i does not know whether the product is suitable for his needs ($\omega_i = 1$) or not ($\omega_i = 0$). As assumed throughout the paper, in the MLM model, when agent i meets a distributor he becomes aware of the product and learns ω_i. Thus, with probability q the agent is willing to pay 1 for the good. Under "traditional marketing," an agent i who is targeted by the SO becomes aware of the product but does not learn ω_i. Thus, the agent is willing to pay q for the product with probability 1. Instead of commissions, the SO incurs a marketing cost of m per agent he targets. Thus, his profit if he targets n agents is $n(q - m)$. In conclusion, for small values of q, the SO is better off in the MLM model, while for large q, the SO can be better off choosing the traditional marketing channel.

An alternative approach that complements the one presented in this paper is to ask whether consumers are buying the company's product or not. If distributors are paying for licenses but consumers are not buying the product this can be indicative of a fraud. However, in practice, sales data is not always easy to obtain. Moreover, it can be difficult to distinguish between sales to other distributors in the network (who may buy to become eligible for extra commissions) and sales outside the network. Finally, as Proposition 2 shows, even if consumers are buying the product and thus the company is not purely a scam, it can still be the case that the distributors are being exploited.

We shall conclude by discussing two modifications of the baseline setting.

Incentive compatibility

Throughout the paper we assumed that the SO uses 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 $\sum_{\tau=1}^{n} b_\tau \leq \eta$ 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.

Below, we extend the network formation model and show that while 1-level schemes cannot sustain a pyramid scam, 2-level schemes can sustain one, even under the stronger incentive-compatibility constraint. Additional illustrations of this effect are given in the Supplemental Appendix, where we consider deterministic social networks, and in

Section 6, where we consider alternative behavioral models. The common component in all of these cases is that there are agents who overestimate the number of downline distributors they will have more than they overestimate the number of people they will recruit themselves. Under ABEE, this occurs only in social networks where there are *some* agents who have fewer successors/friends than their successors/friends have, which is not the case under the uniform random recursive tree model.

Extension: Recruiting one's friends

An individual who joins a scheme may find it natural to first approach his immediate friends as approaching strangers is perhaps more difficult. Such individuals exhaust their best opportunities to recruit new members soon after they join. In order to roughly approximate this, we modify the network formation model such that each player can meet new entrants only in the first period after he enters the game (i.e., an agent who enters the game in period t can meet new entrants only in period $t+1$), and we interpret these entrants as the agent's friends. Moreover, we assume that in each period $t \in \{1, ..., n\}$, $\mu_t > 0$ new agents enter the game such that agents who enter the game in period t have, in expectation, μ_{t+1}/μ_t friends they can make offers to. As in the baseline model, the agent who meets each entrant is drawn by nature uniformly at random. Observe that in this network formation model, *agents do not necessarily meet fewer agents than their successors*.

We now show that, as in the baseline model of Section 4.2, *1-level schemes cannot sustain a pyramid scam* if $a_1 \leq \phi$. Subsequently, we shall show that 2-level schemes can sustain such a scam when $a_1 + a_2 \leq \phi$.

Set $q = 0$ and consider a profile in which every agent who receives an offer up to period $k < n$ accepts it, every agent who receives an offer in period $k+1$ rejects it, and every distributor who meets an agent in periods $1, ..., k+1$ makes him an offer. Note that no offer is made after period $k+1$. Every symmetric ABEE in which the SO's expected profit is strictly positive has this threshold structure under this network formation process. Under this profile, $\mu_1 + ... + \mu_{k+1}$ offers are made and $\mu_1 + ... + \mu_k$ of them are accepted. Hence, $\beta_1 = \frac{\mu_1+...+\mu_k}{\mu_1+...+\mu_{k+1}}$ and $\beta_2 = 1$ are consistent with this profile. Let $\mu_z/\mu_{z-1} = min\{\mu_2/\mu_1, ..., \mu_{k+1}/\mu_k\}$ and consider an agent i who purchases a license in period $z-1$. He analogy-based expects to sell $\frac{\mu_z}{\mu_{z-1}}\beta_1 \leq \frac{\mu_2+...+\mu_{k+1}}{\mu_1+...+\mu_{k+1}} < 1$ licenses, which, if R is an IC 1-level scheme, will not cover the cost of becoming a distributor. Hence, it is impossible to sustain a pyramid scam by means of a 1-level scheme.

The profile that we have described above can be part of an ABEE of a 2-level scheme's induced game. For example, suppose that there are three periods, $k = 1$, $\mu_1 = 1$, $\mu_2 = 4$, and $\mu_3 = 40$. As we calculated above, $\beta_1 = 0.2$. An agent who enters the game in period 1 analogy-based expects to recruit $\beta_1 \mu_2 = 0.8$ distributors and analogy-based expects that these distributors will recruit $\beta_1^2 \frac{\mu_3}{\mu_1} = 1.6$ distributors. If R pays $a_1 = a_2 = 0.5\phi$, an agent who enters the game in period 1 analogy-based expects a payoff of $1.2\phi - \phi - c - 0.8\hat{c}$. An agent who purchases a license in period 2 analogy-based expects to sell $\beta_1 \frac{\mu_3}{\mu_2} = 2$ licenses and, therefore, he analogy-based expects a payoff of $-c - 2\hat{c}$ and thus he finds it optimal not to purchase a license. Hence, for large ϕ, we have an IC 2-level scheme R (that satisfies $a_1 + a_2 \leq \phi$) where $\pi^{ABEE}(R) > 0$.

Time-contingent compensation

Throughout the analysis we focused on stationary schemes. Such schemes are relatively simple and can potentially conceal the non-stationary nature of the environment from the agents. A natural question is whether the stationarity of the rewards limits the SO's expected payoff. In the Supplemental Appendix we address this question by studying the problem of an SO who can use a sequence of schemes $(R_t)_{t=1}^n$, where an agent who purchases a license in period t is paid according to R_t. First, we show that when agents are fully rational the SO maximizes profits by charging a license *fee that decreases over time*. The decreasing fee enables the SO to perfectly discriminate the distributors on the basis of the time at which they purchase a license. This, in turn, allows him to maximize the social surplus and capture all of it. Second, we show that when agents are analogy-based reasoners, the SO can benefit from charging a license fee that *increases over time*. The increasing fee loosens the incentive-compatibility constraint: when the fee increases over time, the commissions can be higher than the fee in some of the schemes in the sequence and so an agent who purchases a license does not need to recruit many new members to recoup the fee paid. As a result, the SO may be able to sustain a pyramid scam using a sequence of 1-level schemes even in instances in which it would be impossible for him to sustain such a scam using a single IC multilevel scheme.

References

[1] Abreu, D. and Brunnermeier, M. K. (2003): "Bubbles and Crashes," *Econometrica,* 71, 173–204.

[2] Babaioff, M., Dobzinski, S., Oren, S., and Zohar, A. (2012): "On Bitcoin and Red Balloons." In *Proceedings of the 13th ACM Conference on Electronic Commerce,* pp. 56–73.

[3] Bénabou, R. and Tirole, J. (2016): "Mindful Economics: The Production, Consumption, and Value of Beliefs," *Journal of Economic Perspectives,* 30, 141–164.

[4] Bhattacharya, P. and Gastwirth, J. (1984): "Two Probability Models of Pyramid or Chain Letter Schemes Demonstrating the Unreliability of the Claims of Promoters," *Operations Research,* 32, 527–536.

[5] Bianchi, M. and Jehiel, P. (2010): "Bubbles and Crashes with Partially Sophisticated Investors," Mimeo.

[6] Bort, R. (2016): "John Oliver Says Multilevel Marketing Companies Like Herbalife Are Pyramid Schemes," *Newsweek*, November 7.

[7] Brunnermeier, M. K. and Parker, J. A. (2005): "Optimal Expectations, *American Economic Review,* 95, 1092–1118.

[8] Camerer, C., Ho, T.-H., and Chong, J.-K. (2004): "A Cognitive Hierarchy Model of Games," *Quarterly Journal of Economics,* 119, 861–898.

[9] Crawford, V. P., and Iriberri, N. (2007): "Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?," *Econometrica*, 75, 1721–1770.

[10] Crawford, V. P., Kugler, T., Neeman, Z., and Pauzner, A. (2009): "Behaviorally Optimal Auction Design: Examples and Observations," *Journal of the European Economic Association*, 7, 377–387.

[11] Crawford, V. P., Costa-Gomes, M. A., and Iriberri, N. (2013): "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications," *Journal of Economic Literature*, 51, 6–52.

[12] DeLong, B., Shleifer, A., Summers, L., and Waldmann R. (1990): "Positive Feedback Investment Strategies and Destabilizing Rational Speculation," *Journal of Finance*, 45, 379–395.

[13] 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.

[14] Direct Selling Association (2016): "Direct Selling in 2016: An Overview." Retrieved from http://www.dsa.org/benefits/research/factsheets.

[15] Drmota, M. (2009): *Random Trees,* Vienna, Springer.

[16] Eliaz, K. and Spiegler, R. (2006): "Contracting with Diversely Naive Agents," *Review of Economic Studies,* 73, 689–714.

[17] Eliaz, K. and Spiegler, R. (2008): "Consumer Optimism and Price Discrimination," *Theoretical Economics*, 3, 459–497.

[18] Emek, Y., Karidi, R., Tennenholtz, M., and Zohar, A. (2011): "Mechanisms for Multi-Level Marketing." In *Proceedings of the 12th ACM Conference on Electronic Commerce*, pp. 209–218.

[19] Eyster, E. and Piccione, M. (2013): "An Approach to Asset Pricing under Incomplete and Diverse Perceptions," *Econometrica*, 81, 1483–1506.

[20] Eyster, E. and Rabin, M. (2005): "Cursed Equilibrium," *Econometrica*, 73, 1623–1672.

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

[22] Federal Trade Commission (2016a): "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.

[23] Federal Trade Commission (2016b): "Multilevel Marketing." Retrieved from https://www.ftc.gov /tips-advice/business-center/guidance/multilevel-marketing.

[24] Gabaix, X. and Laibson, D. (2006): "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics,* 121, 505–540.

[25] Gastwirth, J. (1977): "A Probability Model of a Pyramid Scheme," *American Statistician,* 31, 79–82.

[26] Grubb, M. D. (2009): "Selling to Overconfident Consumers," *American Economic Review,* 99, 1770–1807.

[27] Grubb, M. D. (2015): "Overconfident Consumers in the Marketplace," *Journal of Economic Perspectives,* 29, 9–36.

[28] Guarino, A. and Jehiel, P. (2013): "Social Learning with Coarse Inference," *American Economic Journal: Microeconomics,* 5, 147–174.

[29] Harrison, M. and Kreps, D. (1978): "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics*, 92, 323–336.

[30] Heidhues, P. and Kőszegi, B. (2010): "Exploiting Naïvete about Self-Control in the Credit Market," *American Economic Review,* 100, 2279–2303.

[31] Herbalife (2019): "U.S. Statement of Average Gross Compensation," retrieved from https://edge.myherbalife.com/vmba/media/7B5C662A-D82A-4C1C-AA42-3F21238AAF54/Web/General/Original/StatementOfAverageGrossCompensation.pdf.

[32] Jehiel, P. (2005): "Analogy-Based Expectation Equilibrium," *Journal of Economic Theory*, 123, 81–104.

[33] Jehiel, P. (2011): "Manipulative Auction Design," *Theoretical Economics*, 6, 185–217.

[34] Jehiel, P. and Koessler, F. (2008): "Revisiting Games of Incomplete Information with Analogy-Based Expectations," *Games and Economic Behavior*, 62, 533–557.

[35] Keep, W. and Vander Nat, P. (2002): "Marketing Fraud: An Approach for Differentiating Multilevel Marketing from Pyramid Schemes," *Journal of Public Policy & Marketing*, 21, 139–151.

[36] Keep, W. and Vander Nat, P. (2014): "Multilevel Marketing and Pyramid Schemes in the United States. An Historical Analysis," *Journal of Historical Research*, 6, 188–210.

[37] Koehn, D. (2001): "Ethical Issues Connected with Multilevel Marketing," *Journal of Business Ethics*, 29, 153–160.

[38] Laibson, D. and Yariv, L. (2007): "Safety in Markets: An Impossibility Theorem for Dutch Books," Working Paper 867, Society for Economic Dynamics.

[39] Mahmoud, H. (1991): "Limiting Distributions for Path Lengths," *Probability in the Engineering and Informational Sciences,* 5, 53–59.

[40] McCrum, D. (2016): "Herbalife: Avoiding the P-Word," *Financial Times,* July 18.

[41] McKown, C. (2017): "Legit Business Opportunity or Pyramid Scheme? Six Signs to Watch out for If You're Presented with a Multilevel Marketing Job," *CNBC,* June 9.

[42] Moinas, S. and Pouget, S. (2013): "The Bubble Game: An Experimental Study of Speculation," *Econometrica,* 81, 1507–1539.

[43] Moyer, L. (2018): "Five Years after Brawl with Icahn, Ackman Exits Losing Bet against Herbalife," *CNBC,* February 28.

[44] Mullainathan, S., Schwartzstein, J., and Shleifer, A. (2008): "Coarse Thinking and Persuasion," *Quarterly Journal of Economics,* 123, 577–619.

[45] "Multi-level Marketing in America: Pharaonic Creations. The Growing Battle over How to Spot a Pyramid Scheme" (2015): *The Economist*, October 24.

[46] Piccione, M. and Rubinstein, A. (2003): "Modeling the Economic Interaction of Agents with Diverse Abilities to Recognize Equilibrium Patterns," *Journal of the European Economic Association,* 1, 212–223.

[47] Parloff, R. (2016): "Herbalife Deal Poses Challenges for the Industry," *Fortune*, July 19.

[48] Pierson, D. (2017): "LuLaRoe has Turned Your Facebook Friends into a Leggings Sales Force," *Los Angeles Times*, December 11.

[49] Rantala, V. (2019): "How Do Investment Ideas Spread through Social Interaction? Evidence from a Ponzi Scheme," *Journal of Finance,* 74, 2349–2389.

[50] 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.

[51] Shiller, R. J. (2015): *Irrational Exuberance,* Princeton, Princeton University Press.

[52] Spiegler, R. (2011): *Bounded Rationality and Industrial Organization*, New York, Oxford University Press.

[53] Steiner, J. and Stewart, C. (2015): "Price Distortions under Coarse Reasoning with Frequent Trade," *Journal of Economic Theory*, 159, 574–595.

[54] Suddath, C. (2018): "Thousands of Women Say LuLaRoe's Legging Empire Is a Scam," *Bloomberg*, April 27.

[55] Tirole, J. (1982): "On the Possibility of Speculation under Rational Expectations," *Econometrica*, 50, 1163–1182.

[56] Truswell, N. (2018): "Hun, This Could be Your Opportunity to Get Rich," *BBC News*, February 14.

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

[58] Wieczner, J. (2017): "Herbalife Paid a 200 Million Fine. Then the FTC Screwed It up," *Fortune*, February 2.

Appendix A: Proofs

Proof of Lemma 1. In expectation, regardless of the players' strategies, an agent who purchases a license in period t meets qv_t agents who are willing to pay $1 \geq \eta$ for 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 qv_t units of the good.

Part 1. For sufficiency, note that selling qv_t units yields a payoff of $qv_t b_1 - c - \phi$ 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 $qv_t b_1 - c - \phi$ conditional on purchasing a license. Sequential rationality implies that $qv_t b_1 - c - \phi \geq 0$. Since v_t is decreasing in t, it holds that $b_1 qv_k \geq c + \phi$ 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 $a_1 + b_2 qv_t \geq \hat{c}$ 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) $qv_t b_1 \geq c + \phi$ and (ii) a distributor is supposed to make an offer to an agent (on and off the equilibrium path). Since $qv_t b_1 \geq c + \phi$, 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 $a_1 + b_2 qv_t - \hat{c}$. Sequential rationality implies that $a_1 + b_2 qv_t - \hat{c} \geq 0$. Since v_t is decreasing in t, $a_1 + qb_2 v_k \geq \hat{c}$ holds for every $k \leq t$.

Proof of Theorem 1. Let R be an IC scheme that charges $\phi > 0$ and $\eta \leq 1$, and denote by σ a profile of strategies that is part of a PBE in which the SO's expected profit is $\pi^{PBE}(R)$. Denote $k_1 = sup\{t | b_1 qv_t \geq \phi + c\}$. By Lemma 1, if $k_1 < 2$ under σ, then at most one agent purchases a license in every PBE of $\Gamma(R)$, and so R is not profit-maximizing. In the remainder of the proof, we assume that $k_1 \geq 2$. If $a_1 + b_2 qv_1 < \hat{c}$, let $k_2 = 1$ and, otherwise, let $k_2 = sup\{t | a_1 + b_2 qv_t \geq \hat{c} \textit{ and } qv_t \geq \phi + c\}$. Since v_t is decreasing in t, Lemma 1 implies that every offer made up to (resp., after) period k_1 under σ is accepted (resp., rejected) and every distributor who meets an agent up to (resp., after) period k_2 recruits (resp., does not recruit) the latter.

To show that R is not profit-maximizing, consider a scheme R' such that $\eta' = \eta$, $a'_1 = \phi' = 0$, $b'_1 qv_{k_1} = c$, $b'_2 qv_{k_2} = \hat{c}$, and $a_2 = a_3 = ... = b_3 = b_4 = ... = 0$. Since v_t is decreasing in t, Lemma 1 implies that the agents' behavior in σ is part of a PBE of

$\Gamma(R')$. Hence, to show that $\pi^{PBE}(R') > \pi^{PBE}(R)$ we only need to show that the SO's expected payoff when σ is played is greater in $\Gamma(R')$ than in $\Gamma(R)$.

Since $\eta = \eta'$, 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 $d(SO, i_t) = 1$ (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

$$-\phi + qv_t b_1 \geq -\phi + qv_t(\frac{c+\phi}{qv_{k_1}}) \geq qv_t \frac{c}{qv_{k_1}} = qv_t b'_1, \tag{4}$$

with the second inequality being weak for $t = k_1$ and strict for $t < k_1$.

If $d(SO, i_t) > 1$ (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

$$a_1 - \phi + qv_t(b_1 + b_2) \geq a_1(1 - \frac{v_t}{v_{k_2}}) - \phi(1 - \frac{v_t}{v_{k_1}}) + \frac{v_t \hat{c}}{v_{k_2}} + \frac{v_t c}{v_{k_1}} \geq qv_t(b'_1 + b'_2). \tag{5}$$

Since R is IC, $a_1 \leq \phi$. As $t \leq k_2 \leq k_1$, (5) holds. We conclude that for every $t \leq k_1$ (resp., $t < k_1$) the expected net transfers based on i_t's recruitment and sales are weakly (resp., strictly) lower in R' under σ. Hence, $\pi^{PBE}(R') > \pi^{PBE}(R)$.

Note that $b'_1 \leq \eta'$ since R is IC and that $b'_1 \leq b_1$. If $b'_2 > \eta$, then, in $\Gamma(R')$, the SO incurs a loss of $b'_1 + b'_2 - \eta$ whenever a sale is made by a distributor i such that $d(SO, i) > 1$. The SO can earn more than $\pi^{PBE}(R')$ by using a scheme R'' that is identical to R' except that $b''_2 = 0$ as distributors never recruit in $\Gamma(R'')$. Clearly, R'' is IC. Thus, if R' is not IC, then R'' is IC and $\pi^{PBE}(R'') > \pi^{PBE}(R)$.

We can conclude that R is not profit-maximizing. This means that a profit-maximizing scheme must charge $\eta > 1$ (which induces a payoff of 0) or $\phi = 0$. Incentive compatibility implies that if $\phi = 0$, then $a_1 = 0$ as well.

Proof of Proposition 1. We prove this result by backward induction. Let R be an IC scheme and consider 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 $t \in \{1, ..., n-1\}$ 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 $q = 0$, he cannot expect to meet any agent who will buy the good at a price $\eta > 0$ regardless of the other agents' strategies. Being sequentially rational, he rejects every offer he receives both on and off the equilibrium path. By induction, no agent purchases a license in our PBE and, since $q = 0$, the SO does not sell the good. Thus, $\pi^{PBE}(R) = 0$.

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 makes 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 $t \in \{1, ..., n-1\}$ and suppose that agents reject every offer they receive in periods $t+1, ..., n$. 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 $t+1, ..., n$. Thus, for every offer that is accepted in periods $1, ..., t$ there are, in expectation, v_t rejected offers in periods $t+1, ..., n$, and so the proportion of accepted offers, β_1, cannot exceed $\frac{1}{1+v_t}$. Hence, conditional on accepting an offer, agent i_t falsely expects a payoff of $\beta_1 v_t(a_1 - \hat{c}) - c - \phi$. 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. Since $q = 0$, no agent ever buys the good. Hence, $\pi^{ABEE}(R) = 0$.

Proof of Theorem 3. The proof consists of three parts. Part 1 shows that if $\pi^{ABEE}(R) > 0$ and R is an IC 2-level scheme, then $a_1 \geq \hat{c}$. Part 2 shows that if there exists an IC 2-level scheme R such that $\pi^{ABEE}(R) > 0$ when there are n agents, then there exists an IC 2-level scheme R' such that $\pi^{ABEE}(R') > 0$ when there are $n' > n$ agents. Part 3 establishes that such a 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 $a_1 < c$. A distributor who recruits agent i_t expects to increase his payoff by $a_1 - \hat{c} + \beta_1\beta_2 v_t$. 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 $t < k_2$ falsely expects a payoff of

$$\sum_{j=t+1}^{k_2} \frac{1}{j}\beta_1[(a_1 - \hat{c}) + \beta_1\beta_2 v_j a_2] - c - \phi \leq \beta_1\beta_2 v_t(v_t - v_{k_2})\beta_1 a_2 - c - \phi. \tag{6}$$

Consider the last period t in which an agent accepts an offer in our ABEE. Each offer accepted in periods $1, ..., t$ 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, $\beta_1 \leq \frac{1}{1+v_t-v_{k_2}}$ and $\beta_2 \leq \frac{1+v_t-v_{k_2}}{1+v_t}$. As R is IC, $a_2 \leq \phi$. We conclude that (6) 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 $\pi^{ABEE}(R) > 0$. 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, $\beta_1 = \frac{1}{1+v_k}$ and $\beta_2 = 1$. Hence, agent i_k falsely expects a payoff of

$$\frac{\sum_{j=k+1}^{n} \frac{1}{j}}{1 + \sum_{j=k+1}^{n} \frac{1}{j}}(a_1 - \hat{c}) + \frac{\sum_{j=k+1}^{n-1} \sum_{j'=j+1}^{n} \frac{1}{jj'}}{(1 + \sum_{j=k+1}^{n} \frac{1}{j})^2} a_2 - c - \phi \geq 0. \tag{7}$$

Note that (7) is increasing in n. Thus, for any $n' > n$, there is a scheme $R^{n'}$ that is identical to R except that $\phi^{n'} > \phi$ such that a profile of strategies in which agents accept every offer up to period k, reject every offer made after period k, and in which the SO makes no offers after period k, is part of an ABEE of $\Gamma(R^{n'})$. If $\pi^{ABEE}(R) > 0$ when there are n agents, then $\pi^{ABEE}(R^{n'}) > 0$ when there are n' agents.

Part 3. 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 (7) for $k = 1$. As the harmonic sum diverges, when n goes to infinity, (7) goes to $a_1 - \hat{c} + 0.5a_2 - c - \phi$. Thus, for a sufficiently large n, we can choose ϕ, $a_1 \in (\hat{c}, \phi]$, $a_2 \leq \phi$ and $b_1 = b_2 = 0$ such that (7) holds in equality and the profile σ is part of an ABEE in which the SO's expected 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 $\pi^{ABEE}(R) > 0$, then (7) must hold for some $k \leq n$. For $a_1 \leq \phi$ and $a_2 \leq \phi$, (7) is smaller than $\frac{v_k}{1+v_k}\phi + \sum_{j=k+1}^{n} \frac{v_j}{j}\phi - \phi - c$, which is strictly negative if $\phi \geq 0$, $n \leq 25$, and $k \leq n$.

We now show that for $n \geq 25$ there exists an IC 3-level scheme R such that $\pi^{ABEE}(R) > 0$. Fix a profile σ as described in parts 2 and 3 of the proof of Theorem 3 and recall that σ is consistent with $\beta_1 = \frac{1}{1+v_1}$ and $\beta_2 = 1$ 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

$$\beta_1 \sum_{j=2}^{n} \frac{(x\phi - \hat{c})}{j} + \beta_1^2 \beta_2 \sum_{j=2}^{n-1} \sum_{j'=j+1}^{n} \frac{x\phi}{jj'} + \beta_1^3 \beta_2^2 \sum_{j=2}^{n-2} \sum_{j'=j+1}^{n-1} \sum_{j''=j'+1}^{n} \frac{x\phi}{jj'j''} - \phi - c \qquad (8)$$

and, conditional on purchasing a license, every agent who enters the game after period 1 obtains less than (8). For a sufficiently large n (in particular, for $n \geq 25$), it is possible to choose a large ϕ and an $x < 1$ such that (8) equals 0, in which case, σ is part of an ABEE of $\Gamma(R)$ in which the SO makes a strictly positive expected payoff.

Proof of Proposition 2. In Theorem 3, it was shown that for $n > n^\star$ and $q = 0$, there exists a reward scheme R such that $\pi^{ABEE}(R) = \phi - \hat{c} >> 0$. The fact that $q > 0$ does not change the optimality of the agents' strategies w.r.t. their analogy-based expectations in $\Gamma(R)$ if we set $b_1 = b_2 = 0$. Thus, as long as $B \geq \phi$, $\phi - \hat{c}$ is a lower bound for the SO's expected payoff in a profit-maximizing scheme.

Consider a scheme R' such that $\phi' = 0$. Note that the SO's expected payoff in an ABEE of $\Gamma(R')$ cannot exceed qn, which is his expected revenue in an ABEE in which all agents purchase licenses. Clearly, for a sufficiently small q, $qn < \phi - \hat{c}$.

Consider an IC scheme $\tilde{R}$ such that $\tilde{\phi} > 0$, $\tilde{a}_1 \geq 0$, and $a_2 = a_3 = ... = 0$, and an ABEE (σ, β) of $\Gamma(\tilde{R})$. If agents do not purchase licenses in this ABEE, then, again, the SO's expected payoff cannot exceed qn. Assume that agents do purchase licenses and denote by k_1 the last period in which an agent does so.

If $\tilde{a}_1 \geq \hat{c}$, then, conditional on purchasing a license, every agent makes an offer to every agent he meets. Hence, as shown in the previous proofs, $\beta_1 \leq \frac{1}{1+v_{k_1}}$. Since $\tilde{R}$ is IC, agent i_{k_1}'s analogy-based expected payoff in this case cannot exceed

$$\tilde{\phi} \frac{v_{k_1}}{1 + v_{k_1}} + q \frac{n - k_1}{k_1 + 1} - \tilde{\phi} - c, \qquad (9)$$

where $\frac{n-k_1}{k_1+1}$ is the expected number of agents in the subtree of G rooted at i_{k_1}. If q is sufficiently small, then (9) is strictly negative, in contradiction to the optimality of

i_{k_1}'s strategy.

To complete the proof, let $a_1 < \hat{c}$. Conditional on purchasing a license, agent i_t's expected payoff cannot exceed $q\frac{n-k_1}{k+1+1} - \tilde{\phi} - c < 0$, in contradiction to the existence of an ABEE in which he purchases a license if q is sufficiently small. In conclusion, for a sufficiently small q, the SO's expected payoff given a scheme that charges $\phi = 0$ and/or pays $a_2 = a_3 = ... = 0$ is smaller than $\phi - \hat{c}$, the lower bound for his expected payoff in a profit-maximizing scheme.

Proof of Proposition 3. 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, β^n) be an ABEE of $\Gamma(R^n)$ that induces an expected profit of $\pi^{ABEE}(R^n)$, where $\beta^n = (\beta_1^n, \beta_2^n)$. 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. Without loss of generality, we restrict attention to schemes that charge $\eta^n = 1$.

Throughout the proof, we use several technical results on random trees and perform several related calculations, all of which can be found in the Supplemental Appendix.

Step 1: Lower bounds. There exists a constant $\gamma < 1$ and an integer n' such that, for every $n > n'$, it holds that $\pi^{ABEE}(R^n) > \gamma n$, $k_1^n > \gamma n$, and $k_2^n > \gamma n$. This step is proven in Technical Lemmata 3 and 4 in the Supplemental Appendix.

Step 2: The dual problem. The profit-maximizing scheme must minimize the SO's expected net 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

$$\begin{aligned} & min_{a_1,...,a_{\tau^\star},b_1,...,b_{\tau^\star}} \textstyle\sum_{\tau=1}^{\tau^\star}(a_\tau \kappa(a_\tau) + b_\tau \kappa(b_\tau)) && (10) \\ s.t.\ (i_{k_1}) \quad & \textstyle\sum_{\tau=1}^{\tau^\star}[a_\tau w(a_\tau) + b_\tau w(b_\tau)] - \hat{c}w(a_1) \geq c + \phi \\ (i_{k_2}) \quad & \textstyle\sum_{\tau=1}^{\tau^\star}[a_\tau \hat{w}(a_\tau) + b_\tau \hat{w}(b_\tau)] \geq \hat{c} \\ (IC) \quad & a_1, a_2, ... \leq \phi^n \ and\ b_1, b_2, ... \leq 1, \end{aligned}$$

where $w(z)$ is the marginal increase in i_{k_1}'s willingness to pay for a license due to the commission $z \in \{a_1, b_1, ..., a_{\tau^\star}, b_{\tau^\star}\}$, $\hat{w}(z)$ is the corresponding increase in the distributors' benefit from recruiting agent i_{k_2}, and $\kappa(z)$ is the increase in the SO's cost that is associated with z.

The SO's costs. The SO pays a_τ for every distributor j such that $d(SO, j) > \tau$.

The expected number of such distributors is

$$\kappa(a_\tau) = \sum_{G' \in \mathcal{G}^n} \sum_{j \in G'} \frac{\mathbb{1}\left(d\left(SO, j\right) > \tau\right)}{min\{k_1^n, k_2^n\}!}, \tag{11}$$

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

Constraint $\mathbf{i_{k_1}}$. The LHS is agent $i_{k_1^n}$'s willingness to pay for a license. If $k_1^n \geq k_2^n$, then $w(b_1) = qv_{k_1^n}$ and $w(z) = 0$ for any other commission $z \in \{a_1, ..., a_{\tau^\star}, b_2, ..., b_{\tau^\star}\}$ 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 $w(a_\tau) = (\beta_1^n)^\tau (\beta_2^n)^{\tau-1} \sum_{j=k_1^n+1}^{k_2^n} \frac{l_{j,\tau-1}}{j}$ and $w(b_\tau) = q(\beta_1^n)^{\tau-1}(\beta_2^n)^{\tau-2} \sum_{j=k_1^n+1}^{k_2^n} \frac{l_{j,\tau-1}}{j}$, where $l_{j,\tau}$ is the expected number of agents in the τ-th level of the subtree of G rooted at the j-th entrant. For $\tau = 1$, $w(b_1) = qv_{k_1}$ and $w(a_1) = \beta_1^n \sum_{j=k_1^n+1}^{k_2^n} \frac{1}{j}$. Note that $\hat{c}w(a_1)$ is $i_{k_1^n}$'s expected cost of training new recruits.

Constraint $\mathbf{i_{k_2}}$. The LHS is the increase in 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 $\hat{w}(a_\tau) = (\beta_1\beta_2)^{\tau-1} l_{k_2^n,\tau-1}$ and $\hat{w}(b_\tau) = q(\beta_1\beta_2)^{\tau-2} l_{k_2^n,\tau-1}$.

Step 3: The profit-maximizing scheme is a 2-level scheme. In the Supplemental Appendix (see inequalities (23), (24), and (25)) we use results on random trees to show that, if n is sufficiently large, then, for any $z \in \{a_1, b_2\}$ and $z' \in \{a_3, b_3, ..., a_{\tau^\star}, b_{\tau^\star}\}$, it holds that $\frac{w(z)}{\kappa(z)} \geq \frac{w(z')}{\kappa(z')}$ and $\frac{\hat{w}(z)}{\kappa(z)} \geq \frac{\hat{w}(z')}{\kappa(z')}$ with at least one strict inequality. Because of the linearity of (10) in w, $\hat{w}$, and κ, it follows that if $z' > 0$ and $z' \in \{a_3, a_4, ..., a_{\tau^\star}\}$, then $a_1 = \phi$ and $b_2 = 1$. If this is the case, then the SO's expected profit cannot exceed $\sum_{t=1}^{n} \frac{B+1}{t} + \sum_{t=1}^{n-1} \sum_{j=t+1}^{n} \frac{1}{tj}$, as the SO earns (at most) $1 + B$ from every agent he meets and 1 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.

Step 4: For a sufficiently large n, it holds that $\phi = 0$ at the optimum. Consider an IC scheme R that charges $\phi > 0$. There are two cases to consider, (1) $a_1 \geq c$ and (2) $a_1 < c$. Consider the first case 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, $\hat{w}(a_2) = \hat{w}(a_1) = \hat{w}(b_1) = 0 < \hat{w}(a_1) = 1$. As shown in the Supplemental Appendix (inequality (23)), $\frac{w(a_1)}{\kappa(a_1)} > \frac{w(b_1)}{\kappa(b_1)} > \frac{w(a_2)}{\kappa(a_2)} > \frac{w(b_2)}{\kappa(b_2)}$. Hence, if $a_2 > 0$ or $b_2 > 0$, then $a_1 = \phi$ and $b_1 = 1$, which implies that the SO's expected profit cannot exceed

$\sum_{t=1}^{n} \frac{B+q}{t}$ (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 our scheme is profit-maximizing, then $a_2 = b_2 = 0$.

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

We now show that R is not profit-maximizing by introducing a different scheme, R', such that $\pi^{ABEE}(R') > \pi^{ABEE}(R)$. Let R' be a 2-level scheme such that $a'_1 = a'_2 = \phi' = 0$, $b'_1 = \frac{c}{v_{k_1^n}}$, and $b'_2 = \frac{\hat{c}}{qv_{k_1^n}}$. It is possible to verify that there exists an ABEE of $\Gamma(R)$ in which agents (resp., distributors) accept (resp., make) offers in periods $1, ..., k_1^n$ and reject (resp., do not make) offers in periods $k_2^n + 1, ..., n$. The SO's revenue from the agents' sales in this ABEE is identical to his revenue from the agents' sales in the ABEE of $\Gamma(R)$ 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 are higher than under R'.

In the transition from R to R' the SO's profit is lowered by ϕk_1 since there is no fee under R'. The change from b_1 to b'_1 increases the SO's expected payoff by $\frac{\phi+(a_1-\hat{c})w(a_1)}{qv_{k_1^n}}\kappa(b_1)$, 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 $\frac{\hat{c}}{qv_{k_1^n}}\kappa(b_2)$. Overall, the SO's expected payoff increases in the transition from R to R' if

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

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

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

By incentive compatibility, $\phi \geq a_1$ and, by assumption, $a_1 \geq \hat{c}$. Since $k_1^n > \gamma n$ for large n, it follows that $v_{k_1^n}$ is bounded from above for all n. Note that $k_1^n - \kappa(a_1)$ is equal to the k_1^n-th harmonic number, $h_{k_1^n}$ as it is equal to the expected number of nodes in first level of a random recursive tree. Note that $\kappa(a_1) - \kappa(a_2) > \frac{(h_{k_1^n})^2}{2} - 2$ as it is equal to the expected number of nodes in the second level of a random recursive tree. Thus, for a sufficiently large n the above inequality holds, and so R is not profit-maximizing. We can conclude that a scheme is profit-maximizing and IC only if it charges $\phi = 0$.

Incentive compatibility implies that $a_1 = a_2 = ... = 0$ as well.

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

Proof of Proposition 4. Assume that agents are fully rational. As we showed in Theorem 1, there are three possible cases. In the first, there exists an IC profit-maximizing 2-level scheme $\tilde{R}$ in which $\tilde{\eta} = 1$, $\tilde{a}_1 = \tilde{a}_2 = \tilde{\phi} = 0$, $\tilde{b}_1 = \frac{c}{qv_{k_1}}$, and $\tilde{b}_2 = \frac{\hat{c}}{v_{k_2}}$, where $k_1 \in \{1, ..., n\}$ and $k_2 \in \{1, ..., k_1\}$. In the second, there exists an IC profit-maximizing 1-level scheme $\tilde{R}$ in which $\tilde{\eta} = 1$, $\tilde{a}_1 = \tilde{\phi} = 0$, $\tilde{b}_1 = \frac{c}{qv_{k_1}}$, and $k_1 \in \{1, ..., n\}$. In the third, the SO does not recruit at the optimum and the proof is immediate.

Consider the first case. By Lemma 1, in a PBE of $\Gamma(\tilde{R})$, every agent who receives an offer up to period k_1 accepts it and every agent who receives an offer afterward rejects it. Moreover, every distributor who meets an agent up to period k_2 makes an offer to the latter and every distributor who meets an agent after period k_2 does not make him an offer. We now show that this behavior is part of an ABEE of $\Gamma(\tilde{R})$, which implies that $\pi^{ABEE}(\tilde{R}) \geq \pi^{PBE}(R)$.

In an ABEE of $\Gamma(\tilde{R})$, a distributor who recruits agent i_t increases his (analogy-based) expected payoff by $qv_t\tilde{b}_2 - \hat{c}$. Hence, every distributor who meets an agent up to period k_2 makes an offer to the latter and every distributor who meets an agent after period k_2 does not make him an offer. As a result, an agent who purchases a license in period $t \geq k_2$ does not expect to make offers, and so, conditional on purchasing a license, he analogy-based expects a payoff of $qv_t\tilde{b}_1 - c$. Hence, i_{k_1} is indifferent whether to purchase a license or not. Since v_t is decreasing in t, agents who enter the game after i_{k_1} reject every offer in our ABEE, and, because it is possible to refrain from making offers, agents who enter the game prior to k_1 accept every offer.

The proof of the second case follows the same logic and is omitted for brevity.

Proof of Proposition 5. In a symmetric equilibrium, either all of the agents accept every offer or none of the agents accept any offer. To see this, note that an agent i accepts an offer only if there is at least one agent $j \in I - \{i\}$ such that $\sigma_j(i) = 1$. By symmetry, $\sigma_j(i) = 1$ implies that $\sigma_k(i) = 1$ for every $k \in I - \{i, j\}$. Finally, since the likelihood of meeting new entrants goes down over time, if agents accept every offer they receive from other agents, $p_t(s_{-i}, j)$ is decreasing in t for every $i \in I$ and $j \in I \cup \{SO\}$. As a result, $\sum_{t=1}^{n} PR_i(t|SO)p_t(s_{-i}, SO) \geq \sum_{t=1}^{n} PR_i(t|j)p_t(s_{-i}, j)$ for every $i \in I$ and $j \in I - \{i\}$. Hence, $\sigma_j(i) = 1$ implies that $\sigma_j(SO) = 1$.

In an equilibrium $(s_i)_{i \in I}$ in which agents accept offers, it must be that $r_i(2|j) = 1$ and $r_i(1|SO) = 1$ for every $i \in I$ and $j \in I - \{i\}$. To see this, note that since $p_t(s_{-i}, j)$ is decreasing in t, these beliefs maximize (2) conditional on purchasing a license and the requirement that $r_i(t|j) = 0$ whenever $PR_i(t|j) = 0$. Furthermore, under these beliefs, (3) is greater than (2) such that the corresponding requirement is satisfied. We can conclude that an equilibrium in which agents accept offers exists only if (2) is positive for every $i \in I$ and $j \in I \cup \{SO\}$ given subjective beliefs $r_i(2|j) = 1$ and $r_i(1|SO) = 1$. Since $p_t(s_{-i}, j)$ is decreasing in t, an equilibrium in which agents accept offers exists if and only if (2) is positive when an agent receives an offer from another agent $j \in I$ and all agents accept every offer, i.e.,

$$(1-\psi)\sum_{t=1}^{n} PR_i(t|j)p_t(s_{-i}, j) + \psi p_2(s_{-i}, j) - \phi - c \geq 0. \tag{12}$$

Consider an IC 1-level scheme R. If all of the other agents accept every offer, $p_t(s_{-i}, j) = v_t a_1$. Thus, $\pi^{BP}(R) > 0$ if and only if

$$(1-\psi)\sum_{t=1}^{n} \frac{t-1}{t(n-1-v_1)} v_t a_1 + \psi v_2 a_1 - \phi - c \geq 0. \tag{13}$$

Since the harmonic sum diverges and (13) is continuous and increasing in ψ, if n is sufficiently large, then there is a cutoff $\overline{\psi} < 1$ such that an equilibrium in which agents accept offers exists if and only if $\psi \geq \overline{\psi}$. Hence, a cutoff $\psi^{\star\star}(n)$ exists.

We now show that if n is sufficiently large and $\psi \geq \psi^{\star\star}(n)$, then 1-level schemes are not profit-maximizing. Let R be an IC 1-level scheme such that $\pi^{BP}(R) > 0$. Let R' be a 2-level scheme in which $\phi' = \phi$ and $a'_1 = a'_2 = 0.5a_1$. Under both R' and R, when an agent j is recruited the SO pays a_1 if $d(SO, j) > 2$. If $d(SO, j) = 2$, then under R', the SO pays $0.5a_1$, while under R the SO pays a_1. Hence, if there is an equilibrium of $\Gamma(R')$ in which agents purchase licenses, then $\pi^{BP}(R') > \pi^{BP}(R)$. Such an equilibrium exists if and only if

$$\frac{(1-\psi)a_1}{2}\sum_{t=1}^{n-1} \frac{t-1}{t(n-1-v_1)}(v_t + \sum_{k=t+1}^{n} \frac{v_k}{k}) + \frac{\psi a_1}{2}[v_2 + \sum_{t=3}^{n-1} \frac{v_t}{t}] - \phi - c \geq 0. \tag{14}$$

If $a_1 \leq \phi$ and (13) is positive, then (14) is strictly greater than (13). This has two implications. First, 1-level schemes are never profit-maximizing. Second, since (14) is continuous in ψ, there exists $\psi < \psi^{\star\star}(n)$ such that the SO can sustain a pyramid scam by means of a 2-level scheme. The existence of a cutoff $\psi^{\star}(n) < \psi^{\star\star}(n)$ such that the SO can sustain a pyramid scam if and only if $\phi \geq \psi^{\star}(n)$ follows from (12) being increasing and continuous in ψ given optimal beliefs.

Proof of Proposition 6. Consider an IC 1-level scheme R. An agent of type l_1 believes that all other agents are of type l_0 and expects them to accept each offer with probability α_0. Since R is IC, $a_1 \leq \phi$. Note that $\phi \leq B$ and $\alpha_0 \leq \frac{\hat{c}}{B}$. Thus, the l_1 agent believes that if he purchases a license, he will pay $c + \phi$ and earn an expected net payoff of $a_1\alpha_0 - \hat{c} \leq 0$ from every offer he makes in the future. Clearly, the agent does not find it optimal to purchase a license. To complete the first part of the proof, note that if agents of type l_z do not purchase licenses, then agents of type l_{z+1} do not believe that they can sell licenses and, therefore, they do not purchase licenses either.

Consider an IC 2-level scheme R in which $a_1 = a_2 = 0.5\phi = 0.5B$. Agents of type l_1 who purchase a license in period t and make offers to every agent they meet up to period $k > t$ expect a payoff of

$$\sum_{j=t+1}^{k} \frac{1}{j}(\alpha_0 a_1 + \alpha_0 a_2 \sum_{j'=j+1}^{n} \frac{\alpha_1\alpha_0}{j'} - \hat{c}) - c - \phi. \tag{15}$$

Since the harmonic sum diverges, if n is sufficiently large, there is a period $\tau_1(n)$ such that all agents of type l_1 accept every offer made in periods $1, ..., \tau_1(n)$ and reject every offer made afterward. Moreover, there is a period $t > \tau_1(n)$ such that every l_1 distributor makes an offer to every agent he meets up to period t. Furthermore, (15) implies that $\tau_1(n)$ is monotone in n and goes to infinity when n goes to infinity.

Now suppose that all agents of type l_z accept every offer in periods $1, ..., \tau_z(n)$ and make an offer to every agent they meet in these periods. Moreover, suppose that $\tau_z(n)$ is monotone in n and goes to infinity when n goes to infinity. The argument above (replacing n with $\tau_z(n)$, α_0 with 1, and α_1 with 1) shows that if n is sufficiently large, then there is a period $\tau_{z+1}(n) < \tau_z(n)$ such that all agents of type l_{z+1} accept every offer in periods $1, ..., \tau_{z+1}(n)$ and, conditional on accepting an offer, make an offer to every agent they meet in these periods. Moreover, $\tau_{z+1}(n)$ is monotone in $\tau_z(n)$ and goes to infinity when $\tau_z(n)$ goes to infinity. Hence, for a sufficiently large n, agents of all cognitive types purchase licenses in $\Gamma(R)$.

Supplemental Appendix

This appendix includes three sections. In Section B.1, we study different models of social networks and show that the main results of the paper hold. We start with a model in which the network is deterministic and then study a network formation model with an infinite horizon. In Section B.2 we provide additional results that support several claims made in the text. In Section B.3 we provide technical results that are required for the proof of Proposition 3 and complete its proof.

B.1 Modifying the Social Network

B.1.1 Deterministic social networks

Throughout the analysis, we assumed that each agent is connected to a single agent upstream and that agents are symmetric in their chances of meeting new entrants. A natural question is whether the key insights and results apply when the social network is not a tree and individuals know how many friends they have. In this section, we relax the above assumptions and revisit the setting of Section 4.2, in which we studied pure pyramid scams ($q = 0$).

We examine a model in which the network is deterministic and interpret links as friendships. It is assumed that agents can recruit (only) their direct friends. We first analyze an environment in which the network structure is *commonly known* and then discuss the case in which agents know only their direct friends and extrapolate from summary statistics about the network. In both cases, we shall assume that agents know the probability with which other agents accept offers but neglect the correlation between this probability and the number of friends their friends have (and their position in the network). This correlation neglect leads to two biases. First, agents only partially take into account that their popular friends are likely to have already been sold a distributorship already. Second, agents do not take into account that their more reserved friends are unlikely to find it beneficial to purchase a distributorship.

We start by adapting the baseline model. To simplify the analysis, we set $\hat{c} = 0$. Thus, agents always find it optimal to make offers to all their friends. Let g be a connected social network with $n + 1$ nodes representing the SO and n agents. A link between i and j is denoted by $g_{ij} = 1$ and the absence of such a link by $g_{ij} = 0$. Let $N_j = \{i \in I | g_{ji} = 1\}$ represent j's friends and assume that friendship is reciprocal, that is, $g_{ij} = 1$ if and only if $g_{ji} = 1$.

The timeline in the model is as follows. In period $t = 1$, the SO makes an offer to every agent $j \in N_{SO}$. An agent j who receives an offer in period 1 can accept or reject it. For any $t \geq 1$, let D_t be the set of agents who accepted an offer in period t. In period $t > 1$, every agent $i \in D_{t-1}$ makes an offer to every agent $j \in N_i$ from whom i did not receive an offer up to that point. An agent who receives offers in period t can accept at most one of them or reject all of them, unless he accepted an offer prior to t, in which case he must reject all of the offers received at t. The game ends in the first period t at which no offer is made, i.e., after at most n periods. Agents who purchase a license are paid according to a reward scheme as in the baseline model.

For every period t and agent i, we denote by z_t^i the set of players who made an offer to i in period t and let $\hat{z}_t^i \in \{0, 1\}$ be an indicator that equals 1 if i accepted an offer in period t and 0 otherwise. A private history of length t for agent i is a sequence $(z_1^i, \hat{z}_1^i, z_2^i, \hat{z}_2^i, ..., z_{t-1}^i, \hat{z}_{t-1}^i, z_t^i)$, where $\sum_{j=1}^{t} \hat{z}_t^i \in \{0, 1\}$ and $\cap_{j=1}^{t} z_j^i = \emptyset$. Let H_i^t be the set of length-t private histories that satisfy $\sum_{j=1}^{t-1} \hat{z}_t^i = 0$ and $z_t^i \neq \emptyset$. Let $H_i = \cup_{t=1}^{n} H_t^i$. Agent i's strategy $\sigma_i : H_i \to N_i$ specifies (at most) one offer that i accepts as a function of his private history.

Denote the proportion of accepted offers by $\hat{\beta}$ and observe that

$$\hat{\beta} = \frac{\sum_{t=1}^{n} |D_t|}{N_{SO} + \sum_{t=1}^{n} \sum_{j \in D_t} |N_j - z_1^j - ... - z_t^j|}. \tag{16}$$

Note that the formula for $\hat{\beta}$ may include offers that are rejected for different reasons: offers that are rejected because the agent does not want to purchase a license and offers that are rejected by agents who already bought a license.

We say that an agent's strategy is a best response to $\hat{\beta}$ if the strategy is optimal given a belief that every other agent who receives an offer accepts it with probability $\hat{\beta}$, unless the agent already accepted another offer previously (or accepts another offer made simultaneously), in which case, he rejects the offer. A pair $(\sigma, \hat{\beta})$ is said to form an ABEE if $\hat{\beta}$ equals the proportion of offers accepted to offers made induced by σ, and each agent i's strategy is a best response to $\hat{\beta}$.

The next result establishes that there exists no IC 1-level scheme R such that $\pi^{ABEE}(R) > 0$ regardless of the network structure.

Proposition 7 *For any social network, there exists no IC 1-level scheme R in which $\pi^{ABEE}(R) > 0$.*

Proof. Let g be an arbitrary network. Assume to the contrary that there exists an IC

1-level scheme R such that $\pi^{ABEE}(R) > 0$ and consider an ABEE $(\sigma, \hat{\beta})$ of $\Gamma(R)$. Let $J \neq \emptyset$ be the set of agents who accept offers in this ABEE and, for each $j \in J$, denote by x_j the number of offers j makes in the subsequent period. Let $x^\star = min\{x_j | j \in J\}$ and note that (16) implies that $\hat{\beta} < \frac{1}{x^\star}$. Thus, an agent who makes $x^\star$ offers analogy-based expects to recruit less than $x^\star \frac{1}{x^\star}$ agents, and so analogy-based expects a payoff strictly smaller than $a_1 x^\star \frac{1}{x^\star} - c - \phi$. If R is IC, then $a_1 \leq \phi$, and so our agent finds it strictly suboptimal to purchase a license. Hence, $J = \emptyset$, in contradiction to $J \neq \emptyset$. ■

The next example illustrates that if $n = 7$, then there exist a network structure and an IC reward scheme such that the SO makes a strictly positive expected payoff in their induced game. Note that the network's skeleton is essentially a tree in this example. For larger values of n it is easy to find examples in which a pyramid scam is sustained on a non-tree network. However, this example is chosen since it is relatively simple.

Example 2 *Let $I = \{1, ..., 7\}$ and $g_{SO1} = g_{12} = g_{2i} = 1$ for every $i \in \{3, 4, 5, 6, 7\}$ and assume that there are no additional links. Consider a profile of strategies in which agent 1 accepts the SO's offer and makes an offer to agent 2 who rejects agent 1's offer. Under this profile, $\hat{\beta} = \frac{1}{2}$. Conditional on purchasing a license, agent 1 analogy-based expects a payoff of $\hat{\beta}(a_1 + 5\hat{\beta}a_2) - \phi - c$ and agent 2 analogy-based expects a payoff of $5\hat{\beta}a_1 - \phi - c$. Thus, we can choose $\phi > 0$ so that an IC reward scheme that pays $a_1 = 0$ and $a_2 = \phi$ can support this profile as an ABEE in which the SO makes a profit of ϕ.*

In the above ABEE, agent 1 purchases a license although he has only one friend. If the SO were to use an IC 1-level scheme, the agent would never buy a license because it would be impossible for him to recruit more than one agent. However, given a 2-level scheme, agent 1 may purchase a license since he falsely expects to benefit from the fact that his friend, agent 2, has many friends.

Agent 1's overestimation of the number of downline recruits is greater than his overestimation of the number of people he recruits. In equilibrium, agent 1 does not recruit anyone. However, he falsely expects to recruit 0.5 agents and to have 1.25 downline recruits. This effect occurs when there are agents with significantly more "friends of friends" than friends.

We can embed Example 2 in a larger market to obtain the next result.

Proposition 8 *If $n \geq 7$, then there exists a network and an IC 2-level scheme R where $\pi^{ABEE}(R) > 0$.*

Comment: Extrapolating from partial information about the network structure

The above analysis assumes that agents know the network structure. Let us relax this assumption and assume instead that each agent i knows his friends (i.e., each agent i knows N_i) but does not know the rest of the network. We shall assume that i knows only some summary statistics about the the network.

Before proceeding, it is worth pointing out that, if we maintain the assumption that agents best respond to a belief that their friends accept offers with the average probability $\hat{\beta}$, then the impossibility result of Proposition 7 holds. The reason for this is that the proposition does not rely on any assumption about the network structure or the agents' knowledge of it.

We now assume that each agent i knows the degree distribution in the network. An agent i who naively extrapolates from this distribution may expect to have $|N_i|\hat{deg}$ friends of friends, where $\hat{deg} = \frac{\sum_{i \in I \cup \{SO\}} |N_i|}{n+1}$. Thus, if agent i thinks that every other agent accepts offers with probability $\hat{\beta}$, conditional on purchasing a license after receiving offers from z of his friends, the agent would expect a payoff of

$$\hat{\beta}(|N_i| - z)a_1 + \hat{\beta}^2(|N_i| - z)(\hat{deg} - 1)a_2 - \phi - c \tag{17}$$

in a 2-level scheme. Note that our agent neglects the possibility that other agents already hold a license (with the exception of agents who made him an offer up to that point in the game). The following example establishes that there are network structures that enable the SO to sustain a pyramid scam by means of a 2-level scheme when agents extrapolate naively as described above.

Example 3 *Suppose that there are two cliques of agents, Q_1 and Q_2 (i.e., $g_{ij} = 1$ for every pair of agents $i, j \in Q \in \{Q_1, Q_2\}$). Moreover, suppose that $g_{SO,i} = 1$ for every $i \in Q_1$ and $g_{SO,i} = 0$ for every $i \notin Q_1$. Furthermore, suppose that there are two agents $z_1, z_2 \notin Q_1 \cup Q_2$ such that for every $i \in Q_1$ and $j \in Q_2$ it holds that $g_{z_1 i} = 1$, $g_{z_2 i} = 0$, $g_{z_2 i} = 0$, and $g_{z_2 j} = 1$. Finally, let $g_{z_1 z_2} = 1$. Denote $x = |Q_1|$ and $y = |Q_2|$ and observe that $\hat{deg} = \frac{3x+x^2+2+y+y^2}{3+x+y}$.*

Consider a profile of strategies in which all of the offers made by the SO are accepted and, in period 2, every offer is rejected. Under this profile, $\hat{\beta} = \frac{1}{1+x}$. Thus, (17) becomes $\frac{x}{x+1}a_1 + \frac{x}{(x+1)^2}\frac{3x+x^2+2+y+y^2}{3+x+y}a_2 - \phi$ for every $j \in Q_1$ and $\frac{1}{x+1}a_1 + \frac{1}{(x+1)^2}\frac{3x+x^2+2+y+y^2}{3+x+y}a_2 - c - \phi$ for z_1. Clearly, the former expression is strictly greater than the latter one. Let $a_1 = a_2 = m\phi$. For every x there exists a range of y values and $m < 0.5$ values such that *the former expression is positive and the latter one is negative and, as a result, the profile described above is an equilibrium. Note that in this equilibrium, the SO makes an expected payoff of* $x\phi > 0$.

In the equilibrium that is described above, each $j \in Q_1$ neglects the possibility that his friends have already received an offer. As a result, the agent expects to recruit some of them, failing to realize that it is too late to recruit the rest of the members of Q_1 and that all of these members are going to compete to recruit agent z_1.

Like agent 1 in Example 2, the members of Q_1 believe that they have more friends of friends than friends. In an ABEE in which they accept offers, they expect to recruit no more than one person. Still they find it beneficial to purchase a license as they believe that the people they will recruit are likely to recruit many new members (when y is large).

B.1.2 Infinite horizon

Our main objective in this section is to show that the paper's main insights do not depend on the finiteness of the game. We shall focus on Theorems 1–3 and show that similar results hold when there is uncertainty about the length of the game.

Let us relax the assumption that the game has a fixed number of periods and assume instead that, for each period $t \in \mathbb{N}$, conditional on the game reaching period $t \in \mathbb{N}$, there is a probability of $\delta < 1$ that the game continues and a probability of $1 - \delta$ that it terminates in period t. Note that we can no longer assume that the set of agents is finite. We shall assume that the set of potential entrants is $I = [0, 1]$ and that, in each period $t \in \mathbb{N}$, nature draws one agent $i \in I$ to enter the game uniformly at random. In order to ease the exposition, we shall assume that each agent i's strategy $\sigma_i : \mathbb{N} \to \{0, 1\} \times \{0, 1\}$ is a mapping from time to two decisions: whether or not to purchase a license and whether or not to make an offer.

For each $t \in \mathbb{N}$, the average probability that agents accept an offer in period t is $\bar{\sigma}_t := \int_{j \in I} \sigma_j(t)\, dj$, where $\sigma_j(t) = 0$ (resp., $\sigma_j(t) = 1$) if j rejects (resp., accepts) offers he receives in period t. Let $r_\sigma(t)$ be the objective probability that the t-th entrant receives an offer to purchase a license given the profile σ. We shall say that β_1 is *consistent* with σ if $\beta_1 = \sum_{t=1}^{\infty} r_\sigma(t)\, \bar{\sigma}_t$ whenever $r_\sigma(t) > 0$ for some $t \in \mathbb{N}$. The consistency of β_2 is defined in an analogous manner. As in the main text, an ABEE is a pair of profiles (σ, β) such that the agents' analogy-based expectations, β_1 and β_2, are consistent with σ and each agent's strategy is optimal with respect to β_1 and β_2. The rest of the modeling assumptions remain as in the main text.

Propositions 9–11 are analogous to Theorems 1–3. The proofs of these results are similar to the proofs of the results in the main text except for two main differences. First, the expression for i_t's expected number of direct successors, v_t, changes to

$$v_t = \sum_{j=1}^{\infty} \frac{\delta^j}{j+t}. \tag{18}$$

Second, since the number of periods is not finite, we need to show that there is a period $t^\star \in \mathbb{N}$ such that in every game that is induced by an IC scheme, from period $t^\star$ onward, rejecting every offer to purchase a license is the unique best response of each agent $i \in I$ (regardless of his beliefs about the other agents' behavior). This technical result will allow us to treat the game as one with a finite number of periods.

Lemma 2 *There exists a period $t^\star$ such that for every $t > t^\star$ and every IC reward scheme R, every agent who receives an offer in period t finds it suboptimal to accept it regardless of his beliefs about the other agents' strategies.*

Proof. Consider an agent i_t. In expectation, he will have

$$S_t = \sum_{j=t+1}^{\infty} \frac{\delta^{j-t}}{j} + \sum_{j=t+1}^{\infty} \sum_{j'=j+1}^{\infty} \frac{\delta^{j'-t}}{jj'} + \sum_{j=t+1}^{\infty} \sum_{j'=j+1}^{\infty} \sum_{j''=j+1}^{\infty} \frac{\delta^{j''-t}}{jj'j''} + ... = \frac{\delta}{(t+1)(1-\delta)}$$

successors in G. In an IC scheme, for each of his successors that purchases a license, i_t obtains a commission of (at most) ϕ, and, for each of his successors that purchases the good, i_t obtains a commission of (at most) 1. Hence, i_t's expected payoff conditional on purchasing a license is bounded from above by $S_t(\phi + 1) - \phi - c$. Clearly, there exists $t^\star$ such that for every $t > t^\star$ it holds that $S_t(\phi + 1) - \phi - c < 0$. ■

Fully rational agents

We start with the benchmark result of Section 3. Again, to avoid trivial cases in which the SO does not recruit distributors at the optimum (in which case, all schemes are profit-maximizing), we assume that δ is large. First, note that the proof of Lemma 1 holds as is (with v_t defined as in (18)). Thus, the commissions $a_2, a_3, ..., b_3, b_4, ...$ have no effect on the agents' equilibrium behavior and using them can only lower the SO's expected payoff. The proof of Theorem 1 holds as well, where the assumption that δ is large guarantees that $k_1 \geq 2$. The following proposition concludes this discussion.

Proposition 9 *Fix $q > 0$. There exists $\delta^\star$ such that if $\delta > \delta^\star$ and R is IC and profit-maximizing, then $\phi = a_1 = 0$.*

Analogy-based reasoners

We now consider the main results of Section 4.2. We start with the impossibility result of Theorem 2.

Proposition 10 *Let $q = 0$. There exists no IC 1-level scheme R such that $\pi^{ABEE}(R) > 0$.*

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

Suppose that $a_1 \geq \hat{c}$ and consider an ABEE of $\Gamma(R)$ in which the SO makes offers. By Lemma 2, there is a period $t^\star$ such that, for every $t > t^\star$, regardless of agent i_t's beliefs about the other agents' strategies, it is suboptimal for him to purchase a license. Consider a period $t \in \{1, ..., t^\star\}$ such that agents reject every offer made after period t. Since $a_1 \geq \hat{c}$, every agent who holds a license at the end of period t makes, in expectation, v_t offers after period t. Thus, for every offer that is accepted in periods $1, ..., t$ there are, in expectation, at least $\delta^t v_t$ rejected offers after period t, and so the proportion of accepted offers, β_1, cannot exceed $\frac{1}{1+\delta^t v_t}$. Hence, conditional on accepting an offer, agent i_t makes an analogy-based expected payoff of no more than $\frac{v_t}{1+\delta^t v_t}(a_1 - \hat{c}) - c - \phi$. Since R is IC, $a_1 \leq \phi$, and so i_t's expected payoff conditional on purchasing a license is strictly negative. Thus, in our ABEE, i_t does not purchase a license. By induction, no agent ever purchases a license in our ABEE. Since $q = 0$ no agent ever buys the good. It follows that $\pi^{ABEE}(R) = 0$. ■

The next result corresponds to the possibility result of Theorem 3.

Proposition 11 *Fix $q = 0$. There exists a number $\delta^\star < 1$ such that for every $\delta > \delta^\star$ there exists an IC 2-level scheme R such that $\pi^{ABEE}(R) > 0$.*

Proof. Consider a profile of strategies σ in which (i) every agent accepts an offer if he receives one in period 1 and rejects every offer he receives after period 1, (ii) every distributor makes an offer to every agent he meets, and (iii) the SO makes an offer only to agents he meets in period 1. The analogy-based expectations that are consistent with this profile are $\beta_1 = \frac{1}{1+v_1}$ and $\beta_2 = 1$. Given this profile, the SO's expected payoff is $\phi - \hat{c}$.

Consider a 2-level scheme R such that $a_1 = a_2 = x\phi > 0$. If $x \leq 1$, it is IC. Given σ, the first entrant's perceived expected payoff is

$$\beta_1 \sum_{i=1}^{\infty} \frac{\delta^i}{1+i}(a_1 - \hat{c}) + \beta_1\beta_2 \sum_{i=1}^{\infty} \sum_{i'=i+1}^{\infty} \frac{\delta^{i'}}{(1+i)(1+i')} a_2 - \phi - c. \tag{19}$$

For a sufficiently large $\delta < 1$, the above expression is arbitrarily close to $a_1 - \hat{c} + 0.5a_2 - \phi - c = 0.5x\phi - c - \hat{c}$. Thus, it is possible to set $x < 1$ and $\phi > \hat{c}$ such that (σ, β) is an ABEE of $\Gamma(R)$ in which the SO makes a strictly positive expected payoff. ■

B.2 Additional Results

B.2.1 The SO uses at least two distributors when $qv_4 > c + \hat{c}$

Proposition 12 *If $v_4 > \frac{c+\hat{c}}{q}$ and R is a profit-maximizing scheme, then in every PBE of its induced game at least two agents purchase a license.*

Proof. If no agent purchases a license in $\Gamma(R)$, then the SO's expected payoff is $q + v_1 q$. If only one agent i_t purchases a license, then the SO's expected payoff is $q + v_1 q + \frac{1}{t}[\phi - \hat{c} + qv_t(1 - b_1)]$. As Lemma 1 shows, $b_1 \geq \frac{c+\phi}{qv_t}$. Thus, the SO's expected payoff cannot exceed $q + 2qv_1 - c - \hat{c}$. Consider a scheme R' that pays $b_1' = \frac{c}{v_2}$, $b_2 = \frac{\hat{c}}{v_2}$, $a_1' = a_2' = ... = b_3' = ... = 0$, and charges $\eta' = 1$ and $\phi' = 0$. By our assumption on c, $\hat{c}$, and q, R' is IC. By Lemma 1, in a PBE of $\Gamma(R')$ the first two entrants purchase a license (and only these agents purchase a license). In this PBE, the SO's expected payoff is $q + 2qv_1 + qv_2 - 2c - 2\hat{c} - 0.5q(\frac{c}{qv_2})$, which is greater than $max\{q + v_1 q, q + 2qv_1 - c - \hat{c}\}$ if and only if

$$qv_2 - c - \hat{c} - c\frac{1}{2v_2} > 0.$$

By our assumption on n, $qv_4 \geq c + \hat{c} > c$. Thus, the above inequality holds if $q(\frac{1}{3} + \frac{1}{4}) - \frac{qv_4}{2v_2} > 0$, which holds for $n \geq 4$. We can conclude that in a PBE in which one agent purchases a license the SO's expected payoff is not maximized regardless of the scheme that is used. Thus, at the optimum, at least two agents purchase a license in a PBE. By Lemma 1, agents behave in the same way in all PBEs. ■

B.2.2 Profit-maximizing schemes are not socially optimal

We start by assuming that agents are fully rational and then turn to the case where agents are analogy-based reasoners. In both cases, if nq is sufficiently large, then profit-maximizing schemes are not socially optimal.

Proposition 13 *Suppose that agents are fully rational. There exists a number n' such that if $nq > n'$, then there exists no IC scheme that is both profit-maximizing and socially optimal.*

Proof. Let $k = max\{t|c + \hat{c} < qv_t\}$ and assume to the contrary that R is IC, socially optimal, and profit-maximizing. Being socially optimal, R must incentivize the first k entrants to purchase a license, and must incentivize distributors to make offers to every agent they meet in periods $\{1, ..., k\}$. By Theorem 1, R being profit-maximizing implies that $\phi = 0 = a_1$. By Lemma 1, R pays $b_1 \geq \frac{c}{qv_k}$ and $b_2 \geq \frac{\hat{c}}{qv_k}$. Moreover, R being profit-maximizing implies that $\eta = 1$, $b_1 = \frac{c}{qv_k}$, and $b_2 = \frac{\hat{c}}{qv_k}$ as, otherwise, the SO could increase η, lower b_1, or lower b_2, respectively, without affecting the agents' equilibrium behavior.

Consider a scheme R' in which $b'_1 = \frac{c}{qv_{k-1}}$, $b'_2 = \frac{\hat{c}}{qv_{k-1}}$, $\eta' = 1$, and $a_1 = a_2 = ... = b_3 = ... = \phi = 0$. Under R', only the first $k-1$ agents purchase a license in equilibrium, and so the SO's expected revenue is lower by qv_k than under R. The SO's expected cost is also lower under R' as the commissions are lower and the number of distributors is lower by 1. The expected reduction in the SO's cost is at least

$$qv_k(b_1 + b_2) + (k-1)qv_{k-1}(\frac{c}{qv_k} - \frac{c}{qv_{k-1}}) = c + \hat{c} + c\frac{(k-1)}{kv_k}.$$

By definition, $qv_{k+1} \leq c + \hat{c}$. Thus, the above expression is greater than qv_k if $\frac{c}{c+\hat{c}} \geq \frac{k}{(k^2-1)}$. Note that k goes to infinity when n goes to infinity. Thus, if n is sufficiently large, the SO's expected profit increases in the transition from R to R'. Moreover, if R is IC, then so is R'. We can conclude that R is not profit-maximizing. ■

Proposition 14 *Suppose that agents are analogy-based reasoners. There exists a number n'' such that if $qn > n''$, then there exists no IC scheme that is both profit-maximizing and socially optimal.*

Proof. Let $k = max\{t|c + \hat{c} < qv_t\}$ and assume to the contrary that R is IC, socially optimal, and profit-maximizing. By Proposition 3, if n is sufficiently large, then R being profit-maximizing implies that $b_3 = b_4 = ... = a_1 = a_2 = ... = \phi = 0$. Being socially optimal, R must incentivize the first k entrants to purchase a license, and must incentivize distributors to make offers to every agent they meet in periods $\{1, ..., k\}$. In such a profile, the k-th entrant does not expect to recruit anyone. Hence, $qv_k \geq c$ and, since $b_3 = b_4 = ... = a_1 = a_2 = ... = \phi = 0$, a distributor who recruits the k-th entrant does not expect to earn any rewards from the k-th entrant's downline sales and downline recruitments. Thus, $qv_k \geq \hat{c}$. Since R is profit-maximizing, $qv_k = \hat{c}$ and $qv_k = c$. We can now repeat the argument in the proof of Proposition 13 to show that there exists a scheme R' that induces a higher expected payoff than R for the SO. ■

Comment: Profit-maximizing schemes vs. socially optimal schemes when q is small

When agents are fully rational and q is small, profit-maximizing schemes and socially optimal schemes coincide. However, when agents are analogy-based reasoners, this is not necessarily the case. As shown in the proof of Proposition 2, when $n > n^{\star}$ and q is small, the profit-maximizing scheme incentivizes some of the agents to purchase a license (in fact, most of the SO's profit comes from the fees distributors pay). From a social perspective, the optimal number of distributors is zero when q is sufficiently small. Thus, when q is small and agents are analogy-based reasoners, profit-maximizing schemes are not socially optimal because they *incentivize too many agents to purchase a license* relative to the social optimum.

B.2.3 Time-contingent compensation

In this section, we examine the implications of making the rewards time-contingent. To this end, suppose that the SO can use a sequence of schemes $(R^t)_{t=1}^n$ such that an agent who purchases a license in period t is paid according to R^t. Denote the commissions and fees in the time-t scheme with a superscript t. To account for the fact that an agent who purchases a license in period t recruits new members *after* period t, we adapt the incentive-compatibility constraint and say that a sequence of schemes is IC if $a^t_\tau \leq \phi^{t'}$

and $b^t_\tau \leq \eta^{t'}$ for every $t' > t$ and $\tau \geq 1$. Note that an IC sequence *may include schemes that are not IC*. This observation will play a major role when we analyze the case where agents are analogy-based reasoners.

B.2.3.1 Fully rational agents

Consider the baseline model of Section 3. Note that agents have correct expectations on the equilibrium path of a PBE and that, being rational, no agent purchases a license expecting a payoff smaller than 0. Thus, the SO's expected payoff in a PBE cannot exceed the total surplus at the social optimum (regardless of the scheme that is used). In particular, unless n is small, in the baseline model of Section 3, the SO's expected payoff is strictly smaller than the total surplus as profit-maximizing schemes are not socially optimal (Proposition 13).

Suppose that $max\{t|qv_t \geq c + \hat{c}\} = k$ such that at the social optimum the first k entrants purchase a license. We now show that there exists an IC sequence of schemes that is *socially optimal*, namely, that in a PBE of its induced game the first k entrants purchase a license and other agents do not purchase a license. We also show that in this PBE, each agent obtains an expected payoff of 0. Thus, the SO's expected payoff in this PBE is equal to the total surplus at the social optimum.

Let $b^t_1 = \frac{c}{qv_k}$, $b^t_2 = \frac{\hat{c}}{qv_k}$, $\eta^t = 1$, and $\phi^t = \frac{v_t - v_k}{qv_k} c + \sum_{j=t+1}^{k} \frac{v_j - v_k}{jv_k} \hat{c}$. We can apply Lemma 1 to show that (i) the first k entrants find it optimal to purchase a license, (ii) distributors find it optimal to recruit the first k entrants, and (iii) distributors do not find it optimal to recruit agents after period k. It is easy to verify that, given this behavior, every agent obtains an expected payoff of 0. Thus, the scheme is socially optimal and the SO captures all the surplus by charging a *license fee that goes down over time.*

B.2.3.2 Analogy-based reasoners

When agents are analogy-based reasoners the SO can benefit from charging *a license fee that goes up over time.* As we shall see, this would allow him to make a strictly positive profit in instances in which it would be impossible to do so when he is restricted to using a single scheme. In particular, we shall show that when $n = 2$, the SO can sustain a pyramid scam by means of an IC sequence of two 1-level schemes. The trick here is that the first of the two schemes is not IC. Nonetheless, the increasing fee makes the sequence IC and enables the SO to use it.

To apply the behavioral model of Section 4.2, we shall assume that the time-contingent nature of the rewards does not change the agents' perception of other agents' behavior, namely, the perception that other agents' strategies are time-invariant.

Suppose that $q = 0$ and $n = 2$. Note that there exists no single IC scheme R in which $\pi^{ABEE}(R) > 0$. The reason for this is that the second entrant cannot find it profitable to purchase a license. Therefore, if i_1 accepts the SO's offer, then $\beta_1 \leq \frac{3}{4}$ and i_1's payoff cannot exceed $\frac{3}{4}(a_1 - \hat{c}) - \phi < 0$.

We now show that there exists an IC sequence of 1-level schemes that enables the SO to make a strictly positive expected profit. To see this, consider a profile of strategies in which the SO makes an offer only in period 1, agent i_1 accepts the SO's offer, and i_1 makes an offer in period 2 if he meets agent i_2. Moreover, suppose that agent i_2 rejects every offer regardless of the identity of the proposer. This profile induces the analogy-based expectations $\beta_1 = \frac{3}{4}$ and $\beta_2 = 1$.

Let $\phi^1 = 2\hat{c}$, $a_1^1 = \frac{8}{3}(c + 2\hat{c}) + \hat{c}$, $\phi^2 = a_1^1$, and $a_2^2 \leq \phi^2$. It is easy to verify that the sequence is IC. Given R^1, $\beta_1 = \frac{3}{4}$, and $\beta_2 = 1$, the first entrant finds it optimal to purchase a license and the SO obtains a strictly positive expected payoff. Moreover, the second entrant finds it suboptimal to purchase a license regardless of the schemes used. Thus, the SO can sustain a pyramid scam by means of two 1-level schemes. Note that the *license fee increases over time*, which enables the SO to offer the first entrant a recruitment commission a_1^1 greater than the fee ϕ^1. That is, while R^1 is not IC, the sequence of schemes is IC.

B.3 Technical Results and Proof of Proposition 3

Technical results: Linear bounds for Step 1 of Proposition 3

Lemma 3 *Fix $\gamma' > 0$ such that $q(log\,(1/\gamma')\,(1 - \frac{c+\hat{c}}{q(log(1/\gamma')-1)}) - \hat{c} > \gamma'$. There exists a number $n(\gamma')$ such that for every $n > n(\gamma')$ it holds that $\pi^{ABEE}(R^n) \geq {\gamma'}^2 n$.*

Proof. Denote $k = \lceil \gamma' n \rceil$. Let R be a 1-level scheme in which $\eta = 1$, $\phi = \hat{c}$, $a_1 = \hat{c}$, and $b_1 = \frac{c+\hat{c}}{qv_k}$. Consider a profile σ in which every agent accepts (resp., rejects) every offer he receives up to (resp., after) period k and every 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 $b_1 < \frac{c+\hat{c}}{q(log(1/\gamma')-1)}$. The SO's expected payoff under σ is greater than $k\big(q[1+v_k]\,(1 - b_1) - \hat{c}\big)$. Hence, for a sufficiently large n, it holds that

$$\pi^{ABEE}(R^n) \geq \pi^{ABEE}(R) \geq \gamma' n \left(q(log\,(1/\gamma')\,(1 - \frac{c+\hat{c}}{q(log(1/\gamma')-1)}) - \hat{c} \right) \geq \gamma'^2 n. \quad (20)$$

■

Lemma 4 *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 $1+B$ from every distributor and 1 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 smaller than

$$k_1^n \left(1 + B + v_{k_1^n}\right) + v_{k_1^n} < k_1^n (2 + B + log\,(n/k_1^n)) + log\,(n) + 1.$$

In an ABEE in which $k_2 < k_1$, the SO obtains an expected revenue smaller than

$$k_2^n \left(1 + B + v_{k_2^n}\right) + \sum_{j=k_2^n+1}^{k_1} \frac{v_j + 1 + B}{j} + v_{k_1^n} < k_2^n (2 + B + log(n/k_2^n)) + (1 + B + log(n))^2.$$

There exists $\bar{\gamma} > 0$ that satisfies $\bar{\gamma} log\,(1/\bar{\gamma}) < min\,\{\gamma^2/2, 1/e\}$ 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 (20) is greater than both of the above upper bounds. ■

We can conclude that there exists a constant $\gamma < 1$ and an integer n' such that, for every $n > n'$, it holds that $\pi^{ABEE}(R^n) > \gamma n$, $k_1^n > \gamma n$, and $k_2^n > \gamma n$. This proves Step 1 in the proof of Proposition 3.

Technical results on random trees: Completion of Steps 3 and 4 in Proposition 3.

Lemma 5 *For every $j \in \{1, ..., n-1\}$ and $\tau > 1$ it holds that $v_j l_{j,\tau-1} \geq 2 l_{j,\tau}$.*

Proof. Note that if $j + \tau > n$, then $l_{j,\tau} = 0$. Observe that $v_j l_{j,1} = 2 l_{j,2} + \sum_{z=j+1}^{n} \frac{1}{z^2}$ 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:

$$v_j \left(\frac{l_{j+1,\tau-1}}{j+1} + \frac{l_{j+2,\tau-1}}{j+2} + ... \right) \geq 2 \left(\frac{l_{j+1,\tau}}{j+1} + \frac{l_{j+2,\tau}}{j+2} + ... \right). \quad (21)$$

We can combine the induction hypothesis with the fact that v_j is weakly decreasing in j to see that (21) holds. ■

Calculating the "cost-benefit" ratios $\frac{w}{\kappa}$ and $\frac{\hat{w}}{\kappa}$ (Step 3)

We need to show that, for a sufficiently large n it holds that $\frac{w(z)}{\kappa(z)} \geq \frac{w(z')}{\kappa(z')}$ and $\frac{\hat{w}(z)}{\kappa(z)} \geq \frac{\hat{w}(z')}{\kappa(z')}$ with at least one strict inequality for every $z \in \{a_1, b_2\}$ and $z' \in \{a_3, ..., a_{\tau^\star}, b_3, ..., b_{\tau^\star}\}$.

First, consider the SO's expected cost $\kappa(a_\tau)$ that is given in (11) and note that this is essentially the expected number of the first $x = min\{k_1^n, k_2^n\}$ entrants whose distance is greater than τ. The next lemma shows 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 6 *For every $\tau \in \mathbb{N}$ it holds that*

$$lim_{x\to\infty} \sum_{G' \in \mathcal{G}^n} \sum_{j \in G} \frac{\mathbb{1}\left(d\left(SO, j\right) > \tau\right)}{xx!} = 1. \tag{22}$$

Proof. The tree G is a uniform random recursive tree rooted at the SO, and the distance $d\left(SO, i_t\right)$ corresponds to the *insertion depth* of the $(t+1)$-th node. Theorem 1 in Mahmoud (1991) establishes that the normalized insertion depth $M_i^\star = \frac{M_i - log(i)}{\sqrt{log(i)}}$ has the limiting distribution $\mathcal{N}\left(0, 1\right)$, i.e., the standard normal distribution. Thus, the proportion of nodes inserted at a distance greater than τ from the root on the LHS of (22) goes to 1 when the size of the random tree, x, goes to infinity (to obtain the LHS of (22), note that the number of trees of size $x+1$ is $x!$). ■

By Lemma 4, $min\{k_1^n, k_2^n\} > \overline{\gamma} n$ for a sufficiently large n. Thus, Lemma 6 implies that $lim_{n\to\infty} \frac{\kappa(a_\tau)}{\kappa(a_{\tau+1})} = 1$ and $lim_{n\to\infty} \frac{\kappa(b_\tau)}{\kappa(b_{\tau+1})} = 1$ for every $\tau \in \{2, ..., \tau^\star - 1\}$. Moreover, $lim_{n\to\infty} \frac{\kappa(a_1)}{\kappa(a_2)} = 1$.

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

Case 1. The analogy-based expectations in this case are $\beta_1 = \frac{1}{1+v_{k_1^n}}$ and $\beta_2^n = 1$. Thus, $\frac{w(a_\tau)}{w(b_\tau)} = \frac{1}{q(1+v_{k_1^n})}$. Recall from the main text that $\kappa(b_\tau) = q\kappa(a_\tau) + q\kappa(a_{\tau-1})v_{k_1^n}$ for $\tau > 1$ and $\kappa(b_1) = q\kappa(a_1) + qk_1^n v_{k_1^n}$. By definition, $k_1^n > \kappa(a_1) > ... > \kappa(a_{\tau^\star})$. Moreover, by Lemma 6, $\frac{\kappa(a_\tau)}{k_1^n}$ goes to 1 when n goes to infinity for every $\tau \in \{1, ..., \tau^\star\}$. Thus, for a sufficiently large n it holds that $\frac{w(a_\tau)}{\kappa(a_\tau)} > \frac{w(b_\tau)}{\kappa(b_\tau)}$. By Lemma 5 it holds that $w(a_\tau) \geq 2w(a_{\tau+1})$ and $w(b_\tau) \geq 2w(b_{\tau+1})$. We can conclude that, for a sufficiently large n, it holds that

$$\frac{w(a_1)}{\kappa(a_1)} > \frac{w(b_1)}{\kappa(b_1)} > \frac{w(a_2)}{\kappa(a_2)} > \frac{w(b_2)}{\kappa(b_2)} > \frac{w(z)}{\kappa(z)} \tag{23}$$

for every $z \in \{a_3, ..., a_{\tau^\star}, b_3, ..., b_{\tau^\star}\}$. Note that $\hat{w}(a_1) = 1 > 0 = w(z)$ for every $z \in \{a_2, ..., a_{\tau^\star}, b_1, ..., b_{\tau^\star}\}$.

Case 2. The analogy-based expectations in this case are $\beta_1 = \frac{1}{1+v_{k_1^n}-v_{k_2^n}}$ and

$$\beta_2^n = \frac{k_1^n - \sum_{t=1}^{k_1^n} \frac{1}{t} + k_1^n(v_{k_1^n} - v_{k_2^n})}{(k_1^n - \sum_{t=1}^{k_1^n} \frac{1}{t} + k_1^n v_{k_1^n})},$$

where $\sum_{t=1}^{k_1^n} \frac{1}{t}$ represents the SO's expected number of offers in periods $t = 1, ..., k_1^n$. Hence, $\beta_1^n \beta_2^n \leq \frac{1}{1+v_{k_1^n}}$. Since $\frac{k_1^n}{k_1^n - \sum_{t=1}^{k_1^n} \frac{1}{t}}$ goes to 1 when n goes to infinity, for a sufficiently large n, $\beta_1^n \beta_2^n$ is arbitrarily close to $\frac{1}{1+v_{k_1^n}}$. Note that $\frac{q}{\beta_1^n \beta_2^n} w(a_\tau) = w(b_\tau)$ for $\tau > 1$. Thus, for a sufficiently large n, $\frac{w(a_\tau)}{\kappa(a_\tau)}$ is arbitrarily close to $\frac{w(b_\tau)}{\kappa(b_\tau)}$ for $\tau > 1$. As above, we can use Lemma 5 to show that $w(a_\tau) \geq 2w(a_{\tau+1})$ for $\tau \geq 1$ and $w(b_\tau) \geq 2w(b_{\tau+1})$ for $\tau > 1$. We can conclude that

$$\frac{w(a_1)}{\kappa(a_1)} > \frac{w(b_2)}{\kappa(b_2)} > \frac{w(z)}{\kappa(z)} \tag{24}$$

for every $z \in \{a_3, ..., a_{\tau^\star}, b_3, ..., b_{\tau^\star}\}$. We can repeat the above exercise for the ratio $\frac{\hat{w}}{\kappa}$ (the proof is identical to the one for the ratio $\frac{w}{\kappa}$ and, therefore, it is omitted).

Case 3. The analogy-based expectations in this case are $\beta_1^n = 1$ and

$$\beta_2^n = \frac{k_2^n - \sum_{t=1}^{k_2^n} \frac{1}{t}}{k_2^n - \sum_{t=1}^{k_2^n} \frac{1}{t} + k_2^n v_{k_2^n} + \sum_{j=k_2^n+1}^{k_1^n} \frac{v_j}{j}}.$$

It is easy to see that $w(b_1) > 0$ and $w(z) = 0$ for every $z \in \{a_1, ..., a_{\tau^\star}, b_2, ..., b_{\tau^\star}\}$, as agent $i_{k_2^n}$ does not expect to sell licenses. It is left to consider the ratio $\frac{\hat{w}}{\kappa}$. Clearly, $\hat{w}(a_1) = 1 > \frac{v_{k_2^n}}{1+v_{k_2^n}} \geq \beta_1^n \beta_2^n v_{k_2^n} = \hat{w}(a_2)$. Since $\beta_1^n \beta_2^n < \frac{1}{1+v_{k_2^n}}$, Lemma 5 implies that $\hat{w}(a_\tau) \geq 2\hat{w}(a_{\tau+1})$ for every $\tau > 1$. Similarly, Lemma 5 implies that $\hat{w}(b_\tau) \geq 2\hat{w}(b_{\tau+1})$ for every $\tau > 1$.

Note that Lemma 4 implies that $k_2^n > \bar{\gamma} n$ for large n. As a result, $v_{k_2^n}$ is bounded from above (e.g., by $log(\frac{1}{\bar{\gamma}})+1$). Furthermore, for a sufficiently large n, β_2^n is arbitrarily close to $\frac{1}{1+v_{k_2^n}}$. This implies that, for a sufficiently large n, $\hat{w}(a_\tau)$ is arbitrarily close to $\hat{w}(b_\tau)$ for any $\tau > 1$. Moreover, since $v_{k_2^n}$ is bounded from above, $\hat{w}(a_2)$ is bounded below $1 = \hat{w}(a_1)$. We can conclude that, for a sufficiently large n, it holds that

$$\frac{\hat{w}(a_1)}{\kappa(a_1)} > \frac{\hat{w}(b_2)}{\kappa(b_2)} > \frac{\hat{w}(z)}{\kappa(z)} \tag{25}$$

for every $z \in \{a_3, ..., a_{\tau^\star}, b_3, ..., b_{\tau^\star}\}$.