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

Yair Antler[†]

November 30, 2018

Abstract

We identify the conditions on the tendency of agents to spread information by word of mouth, under which a principal can design a pyramid scam to exploit a network of agents whose beliefs are coarse. We find that a pyramid scam is sustainable only if its underlying reward scheme compensates the participants on multiple levels of recruitments (e.g., for recruiting new members and for recruitments made by these new members). Motivated by the growing discussion on the legitimacy of multilevel marketing schemes and their resemblance to pyramid scams, we compare the two phenomena based on their underlying compensation structure.

What delineates pyramid scams from legitimate multilevel marketing enterprises? Dramatic 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

*I am particularly grateful to Rani Spiegler for many insightful discussions and useful suggestions. I thank Vellore Arthi, Benjamin Bachi, Kfir Eliaz, Erik Eyster, Michael 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 EARIE 2018, 17th Annual Berkeley/Columbia/Duke/MIT/Northwestern IO Theory Conference, Frankfurt, Hebrew University, LSE, Royal Holloway, Tel Aviv, Toulouse, UCL, and Warwick.

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

[1]Membership in MLMs is substantial and growing. For instance, according to a recent New York Times article, Amway has 1.9 million distributors in China alone (McMorrow and Myers, 2018).

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

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.

The MLM industry's questionable legitimacy and the resemblance of its companies to fraudulent pyramid scams received considerable attention in the mainstream media[3] following a recent feud between Herbalife and the hedge-fund tycoon Bill Ackman, a dispute that led to an FTC investigation against the former party (FTC, 2016a). Identifying whether a particular company is a legitimate one, or whether it is an exploitative pyramid scam that promotes useless goods 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 if it prioritizes recruitment over selling (FTC, 2016b). 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 based on their underlying compensation schemes. The premise of our analysis is that the potential distributors are strategic, and that the MLM company (or the pyramid organizer) chooses a compensation scheme while taking these prospective distributors' incentives into account. To understand the structure of the potential reward schemes, consider the following example.

Example 1 *The reward scheme R pays every distributor a commission of b_1 for every retail sale that he makes and a commission of a_1 for every agent that he recruits to the sales force. The reward scheme R' pays every distributor a commission of b'_1 for every retail sale that he makes and a commission of b'_2 for every retail sale made by one of his recruits. It also pays every distributor a'_1 for every one of his 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 one-level (respectively, multilevel) schemes as they compensate the distributors based on the first level (respectively, multiple levels) of their downline.*

Observe that both R and R' compensate the distributors for recruiting others to work for the company. In practice, however, the bulk of the MLM industry uses multilevel

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

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

schemes (DSA, 2014) such as R', rather than one-level schemes such as R. Moreover, even though there is no obvious reason why one-level schemes such as R cannot be used for the purpose of sustaining a pyramid scam, various companies that were deemed[5] pyramid scams used multilevel reward schemes. What can explain these stylized facts? Can a legitimate company benefit from charging entry 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?

In order to address the above questions, we develop a model of word-of-mouth marketing in which a scheme organizer (SO) tries to sell a good to a network of agents that is formed randomly and sequentially. In order to reach larger parts of the network, the SO sells distribution licenses to some of the agents. Distributors can sell units of the good as well as distribution licenses, and they are compensated according to a reward scheme that is chosen in advance by the SO. 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 this "business opportunity" unattractive to agents who receive an offer to join the sales force late in the game, when there are fewer opportunities to sell the good and recruit others to the sales force.

Assume for a moment that the good has no intrinsic value such that the only "products" that are being traded are distribution licenses. If there exists a reward scheme such that the SO makes a strictly positive expected profit in its induced game, then we have a *pyramid scam*. Pyramid scams are, roughly speaking, zero-sum games and, therefore, fully rational economic agents will never participate in them. Nevertheless, we observe countless such scams in practice (see, e.g., Keep and Vander Nat, 2014, and the references therein). Hence, if we wish to better understand such scams and their underlying compensation schemes, we must depart from the classic rational expectations framework. We shall use Jehiel's (2005) elegant framework of *analogy-based expectation equilibrium* to relax the requirement that the agents have a perfect understanding of the other agents' behavior in every possible contingency, while maintaining that the agents' beliefs are *statistically correct*.[6]

Under the behavioral model, each agent correctly predicts (and best responds to) the other agents' *average behavior*. However, he neglects the fact that the other agents' strategies might be time-contingent. This mistake leads each agent to mispredict the

[5] See, e.g., Fortune High-Tech Marketing (*FTC v. Fortune Hi-Tech Marketing, Inc.*, 2014) and BurnLounge (*FTC v. BurnLounge, Inc.*, 2014).

[6] In Section 4, we shall discuss the implications of other behavioral models for our results.

other agents' "marginal" equilibrium behavior (e.g., to think that it might be possible to recruit new participants to the pyramid late in the game). 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, or as learning from partial feedback about their behavior in similar past interactions (e.g., the sales force's average performance in past schemes organized by the SO).

An individual who contemplates joining a pyramid often tries to assess his ability to recruit others as well as his potential recruits' respective ability. A general insight that emerges from the model is that individuals who understand others' average behavior *do not overestimate their own ability to recruit by much*, if at all. Therefore, the SO cannot exploit such individuals and sustain a pyramid scam by means of one-level schemes such as R. Multilevel schemes such as R' introduce additional variables for the prospective participants to mispredict (e.g., their recruits' ability to recruit). All of these prediction mistakes are small. However, the *accumulation of these small prediction errors* enables the SO to sustain a pyramid scam.

We provide necessary and sufficient conditions on the number of agents and their tendency to spread information by word of mouth under which the SO can sustain a pyramid scam, and we show that multilevel schemes can support such a scam whereas one-level schemes cannot generate a strictly positive expected profit for the SO.

In order to better understand legitimate MLM, we shall examine a setting in which the goods have an intrinsic value such that the SO benefits from selling them. The reward scheme's objective in this case is to incentivize the distributors both to sell the products and to propagate information about them, while maintaining a low overhead. We solve for the SO's optimal scheme under two behavioral assumptions. First, we show that if the agents are fully rational, then schemes that maximize the SO's expected equilibrium profit do not charge license fees, nor do they pay recruitment commissions. Second, when the SO faces a population of analogy-based reasoners, then if the number of agents is sufficiently large, under mild assumptions, optimal schemes compensate the distributors only for sales and do not charge license fees. Thus, the two pyramidal components—recruitment commissions and license fees—are not in use when the good is intrinsically valued even though the agents are vulnerable to deceptive practices.

The main contribution of the paper is fourfold. First, we develop a model that enables us to better understand what makes pyramid scams work. Second, our results suggest an explanation as to why such scams often rely on multilevel reward schemes and dubious "passive income" promises, which are one of the hallmarks of pyramid scams according to the Securities and Exchange Commission's (SEC) investor alert

(SEC, 2013). Third, our analysis shows how the presence of rational agents can prevent vulnerable agents from participating in exploitative pyramid scams in noncompetitive environments. Finally, our analysis of legitimate MLM allows us to more transparently draw the dividing line between exploitative scams and legitimate MLM enterprises.

Related literature

We use analogy-based expectation equilibrium, which was developed in Jehiel (2005) and extended in Jehiel and Koessler (2008), as our behavioral framework. A closely related concept, "cursed equilibrium," was developed by Eyster and Rabin (2005) for games of incomplete information. In a 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 reason in terms of a coarse representation of the world are Mullainathan et al. (2008), Jehiel (2011), Eyster and Piccione (2013), Guarino and Jehiel (2013), and Steiner and Stewart (2015). In Eyster and Rabin (2010), in the context of social learning, agents best respond to a belief that their predecessors are cursed (i.e., do not learn from their own predecessors' behavior).

Our work relates to a strand of the behavioral industrial organization literature in which rational firms exploit boundedly rational agents. 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. Grubb (2009) studies contracting when agents are overconfident about the accuracy of their forecasts of their own future demand. In Laibson and Gabaix (2006), firms may hide information about add-on prices from unaware consumers. Heidhues and Kőszegi (2010) study exploitative credit contracts when consumers are time-inconsistent. In the context of auctions, Crawford et al. (2009) show that agents who are characterized by level-k thinking can be exploited by a rational auctioneer. Eliaz and Spiegler (2007, 2008, 2009) apply a mechanism design approach to speculative trade.

This article is also related to the "Dutch Books" literature that studies the vulnerability of nonstandard preferences or of nonstandard decision-making procedures to exploitative transactions. For example, Rubinstein and Spiegler (2008) examine the extent to which agents who employ a sampling procedure in the spirit of the S-1 equilibrium are vulnerable to exploitative transactions offered by a rational market maker. Laibson and Yariv (2007) show that competitive markets may protect agents with nonstandard preferences from exploitative schemes.

Shiller (2015) describes speculative bubbles as *naturally occurring pyramid scams* (i.e., bubbles do not include the design element). Tirole (1982) shows that such bubbles cannot exist under the classic rational expectations model. Harrison and Kreps (1978) suggest a resale option theory of bubbles that is based on heterogeneous prior beliefs. In their model, investors overpay for an asset hoping to resell it at an even higher price in the future. A different strand of the literature suggests a "feedback loop" theory of bubbles (Shiller, 2015). For example, in DeLong et al. (1990), rational speculators anticipate that positive-feedback noise traders will push an asset's price above its fundamental value in the future, and therefore purchase the asset in order to resell it at an inflated price. These purchases fuel the noise traders' demand and inflate the price further. Bianchi and Jehiel (2010) show that the analogy-based expectation equilibrium logic can sustain both bubbles and crashes in equilibrium. We shall discuss Abreu and Brunnermeier's (2003) work in detail in the concluding section.

Pyramids and MLM have received considerable attention outside of the economics literature. A strand of the computer science literature (e.g., Emek et al., 2011) focuses on MLM mechanisms' robustness to Sybil attacks. Babaioff et al. (2012) study similar mechanisms in the context of Bitcoin. The marketing literature has addressed ethical issues in multilevel marketing and the resemblance of such schemes to pyramid scams (a comprehensive overview is provided in Keep and Vander Nat, 2014). 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 retail sales to end users (see, e.g., Koehn, 2001; Keep and Vander Nat, 2002). Gastwirth (1977) and Bhattacharya and Gastwirth (1984) use the random recursive tree model to examine two real-world scams and demonstrate that only a small fraction of the participants can cover the entry fees. In none of the above-mentioned models, however, is there strategic interaction.

The paper proceeds as follows. We present the model in Section 1 and analyze pyramid scams in Section 2. Section 3 examines legitimate MLM. Section 4 examines the implications of prominent behavioral models for our results and Section 5 concludes. All proofs are relegated to Appendix A. In Appendix B, we present a natural semistationary modification of the model and two technical results.

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 $i \in I$ is characterized by a unit demand and two numbers: his willingness to pay $\omega_i \in \{0, 1\}$ and his talkativeness $\psi_i \in \{0, 1\}$. The term "talkativeness" will be clarified soon. For every agent $i \in I$, we assume that ω_i and ψ_i are drawn independently and we denote $p := Pr(\psi_i = 1)$ and $q := Pr(\omega_i = 1)$, respectively.

Time $t = 1, 2, ..., n$ is discrete. In each period t, nature draws a new agent (uniformly at random) who enters the game and meets one player who is chosen uniformly at random from a group of players that includes the SO and every agent who entered the game prior to period t. For example, the second entrant meets either the SO or the first entrant, each with probability 0.5. We often use i_t to denote the t-th entrant. Let G denote the directed tree, rooted at the SO, that results at the end of this process and let G_i denote the subtree of G rooted at $i \in I$. We denote the length of the directed path between $i \in I \cup \{SO\}$ and $j \in I$ by $d_G(i, j)$.

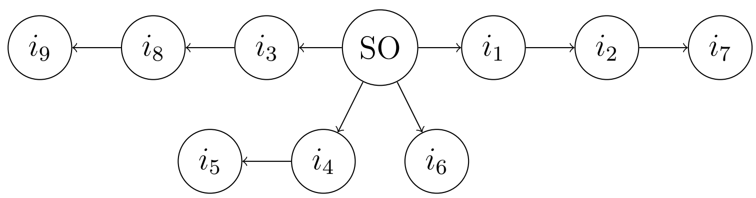

Figure 1: A snapshot of G at the end of period 9.

For each $t \geq 1$, let D_t be a set that includes the SO and the distributors at the beginning of period t. The t-th entrant can become a distributor if he meets a member of D_t. If he meets an agent $i \notin D_t$, he can still become a distributor if every agent on the directed path connecting him to the nearest member of D_t is talkative (i.e., if $\psi_j = 1$ for every agent j who is on that path).

Formally, for each $t \geq 1$, there exists at most one player $j \in D_t$ such that there is a directed path connecting j to the t-th entrant and every agent on that path is (i) talkative and (ii) not a distributor (i.e., $\psi_l = 1$ and $l \notin D_t$ for every agent l on that path). If such a player $j \in D_t$ exists, then, in period t, j can offer the t-th entrant the opportunity to become a distributor.[7] If the t-th entrant receives an offer, then he can accept it and become a distributor or reject it. In addition, regardless of whether an offer is made by j, the t-th entrant purchases a unit of the good for personal consumption from j at a price η^R that is predetermined by the SO if and only if[8] $\omega_{i_t} \geq \eta^R$. If such a player $j \in D_t$ does not exist, then the t-th entrant does not receive an offer to become a distributor and does not purchase the good.

To understand the game, consider a history such that, at the end of period 9, G is as presented in Figure 1, $D_{10} = \{SO, i_1, i_2, i_3\}$, $\psi_{i_4} = 0$, and $\psi_{i_5} = \psi_{i_6} = \psi_{i_7} =$

[7]Property (ii) implies that from the t-th entrant's perspective, j is the nearest member of D_t.

[8]It is possible to make the price and the decision to sell/buy the good endogenous without changing the main results in this article.

$\psi_{i_8} = \psi_{i_9} = 1$ (i.e, i_4 is not talkative and i_5,i_6,i_7,i_8, and i_9 are talkative). Recall that i_{10} is equally likely to meet each of the players who entered the game prior to period 10. If i_{10} meets a player $j \in D_{10}$, then j decides whether or not to make i_{10} an offer to become a distributor. If i_{10} meets a non-distributor $j \in \{i_4, i_5\}$, then i_{10} will not receive an offer or purchase the good as the path from the SO is "blocked" by agent i_4 who is not talkative and not a distributor. However, if i_{10} meets a non-distributor $j \in \{i_6, i_7, i_8, i_9\}$, then j will refer i_{10} to the nearest member of D_{10} who will then decide whether or not to make an offer to i_{10}. Thus, i_7 will refer i_{10} to i_2, i_6 will refer him to the SO, and i_9 will refer him to i_3 via i_8.

Reward schemes, payoffs, and information

The distributors are paid according to a reward scheme that is chosen in advance by the SO. Each reward scheme R includes four components:

- An entry fee $\phi^R \geq 0$.
- Recruitment commissions: $a_1^R, a_2^R, a_3^R, ... \geq 0$.
- Sales commissions: $b_1^R, b_2^R, b_3^R, ... \geq 0$.
- A price $\eta^R \geq 0$ at which each unit of the good is sold.

When the t-th entrant purchases a unit of the good (respectively, becomes a distributor), each distributor $l \in D_t$ who is on the path connecting the SO to the t-th entrant obtains a commission of b_{y+1}^R (respectively, a_{y+1}^R) from the SO, where $y \geq 0$ is the number of distributors on the path connecting l to the t-th entrant. In addition, the SO receives η^R (respectively, ϕ^R) from the t-th entrant.

We assume that each agent who becomes a distributor incurs a cost of $c \geq 0$ that reflects learning about the good and how to sell it. When an agent contemplates purchasing a distribution license (i.e., becoming a distributor), he weighs the expected sum of commissions that he will obtain given R and his beliefs about the other players' behavior against the total cost of becoming a distributor $c + \phi^R$.

The SO faces the risk that the distributors will create fictitious recruits in order to become eligible for additional commissions.[9] Motivated by this risk, we shall focus on schemes where $a_\tau^R \leq \phi^R$ and $b_\tau^R \leq \eta^R$ for each $\tau \geq 1$, and refer to such schemes as *incentive-compatible* (IC) schemes. The IC constraint implies that for a distributor, the cost of creating a fictitious new tree of sales and recruits is greater than the direct

[9]In the computer science literature, a reward scheme's robustness to manipulations in this spirit is often referred to as robustness to local false-name manipulations or robustness to local splits (see, e.g., Emek et al., 2011; Babaioff et al., 2012).

benefit of doing so (i.e., the transfers from the SO to the root). The IC constraint rests on the assumption that the SO can verify[10] 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 would be eligible to receive.

Every scheme R induces a game, which we shall denote by $\Gamma(R)$. The SO's highest expected profit in an equilibrium of $\Gamma(R)$ is denoted by $\pi(R)$. We assume that t, q, p, and the network formation process are commonly known and that, for each $i \in I$, ω_i and ψ_i are i's private information. For each $i \in I \cup \{SO\}$, H_i is the set of nodes in which i must move. Player i's ($i \in I \cup \{SO\}$) strategy is a mapping $\sigma_i : H_i \to \{0,1\}$, where in each $h \in H_i$, i decides whether or not to make an offer, or else decides whether or not to accept one.[11] To simplify the exposition, we shall assume that if an agent $i \in I$ is indifferent between making an offer and not doing so, then he makes it (none of the results in the paper are sensitive to this assumption).

Discussion: Modeling assumptions

Meeting process. We borrow the meeting process from the applied statistics literature, where it is referred to as the *uniform random recursive tree model* (for a textbook treatment, see Drmota, 2009). This process rests on the assumptions that there is a deterministic date at which the game ends and that the number of entrants in each period t is independent of t. 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. To demonstrate the robustness of our results, in Appendix B we modify the process such that conditional on reaching period t, there is a probability $\delta < 1$ that the process continues for an additional period and a probability $1 - \delta$ that the process ends immediately. In addition, in the concluding section, we illustrate the robustness of our results to a more "traditional" branching process in which new entrants can only meet "leaves" in the existing tree.

Word of mouth. The word-of-mouth parameter p adds a realistic aspect to the model, where talkativeness can be interpreted as an agent's tendency to mention the good to others even when he does not have any financial incentive to do so (i.e., when the agent is not himself a distributor). In addition to making the model more realistic, p also provides a natural rationale for recruitment-based compensation in a "legitimate"

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

[11] Our results are robust to any assumption about what a player who moves in period t knows about the events that took place prior to t (e.g., the network's realization, whether offers were made, whether licenses were sold, etc.) and, therefore, we refrain from making any such assumption.

setting where the SO faces fully rational agents. When $p > 0$, a distributor l who sells a distribution license to an agent $j \in I$ loses his direct access to j's successors but improves the chances that j's successors will purchase the good (e.g., directly from j). Thus, in order to incentivize such information propagation, the reward scheme must compensate the distributors for such losses. Surprisingly, as we shall see in the sequel, the talkativeness of non-distributing agents has an additional nontrivial negative effect on the SO's ability to sustain a pyramid scam when he faces boundedly rational agents.

2 Pure Pyramid Scams ($q = 0$)

In order to capture the idea that the only "product" that is being traded in a real-world pyramid scam is the right to recruit others to the pyramid, we set $q = 0$. Thus, it is commonly known that the only products that are being traded in the model are distribution licenses (when an agent accepts an offer to become a distributor, we say that he purchases a distribution license from the distributor who made the offer). Intuitively, such a market should not exist as trade in distribution licenses does not add value. If $q = 0$ and there exists a scheme R such that $\pi(R) > 0$, then we say that the SO is *able to sustain a pyramid scam*. The next result establishes that when all of the agents are fully rational, the SO cannot sustain such a scam.

Proposition 1 *Let $q = 0$. There exists no IC reward scheme R such that the SO makes a strictly positive expected profit in a subgame perfect Nash equilibrium of $\Gamma(R)$.*

Since $q = 0$, reward schemes induce zero-sum transfers between the agents and the SO. Proposition 1 then follows directly 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. We shall weaken the Nash equilibrium assumption that agents have complete understanding of the other agents' behavior in every possible contingency, an assumption that might be too extreme in complicated settings such as the present one.

2.1 The behavioral model

Jehiel (2005) suggests an elegant framework that incorporates partial sophistication into extensive-form games. We adopt this framework and use *analogy-based expectation equilibrium* to solve the model. In an analogy-based expectation equilibrium, different contingencies are bundled into analogy classes and the agents are 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 that the other agents' behavior can be time-contingent. In simple words, agents do not base their expectations that offers are accepted on the time at which they are made. Instead, they consider the average rate of offer acceptances, pooling all offers made at any point in time. Thus, each agent views the other agents' behavior as if it were time-invariant.

In equilibrium, the agents' beliefs about the other agents' behavior are statistically correct. These beliefs 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 that the SO has organized). 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 (e.g., reports that contain summary statistics on the performance of the sales force in past schemes are often publicly available).

For each $i \in I$, denote by H_i^1 (respectively, H_i^2) the set of nodes in which i decides whether or not to purchase a license (respectively, offer a license). Let $M_1 := \cup_{i \in I} H_i^1$ and $M_2 := \cup_{i \in I} H_i^2$. We shall use $\sigma = (\sigma_i)_{i \in I \cup \{SO\}}$ to denote a profile of strategies, and $r_\sigma(h)$ to denote the probability of reaching $h \in M_1 \cup M_2$, conditional on σ being played. For each $h \in M_1 \cup M_2$, we use $\sigma(h) = 1$ (respectively, $\sigma(h) = 0$) to denote that the agent who moves at h makes an offer or accepts one (respectively, makes no offer or rejects one). For each $i \in I$, $\beta^i = (\beta_1^i, \beta_2^i)$ are agent i's analogy-based expectations.

Definition 1 *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 $h \in M_k$.*

A strategy σ_i is a best response to β^i if it is optimal given the belief that every agent $j \neq i$ accepts every offer that 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 *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 to denote the agents' analogy-based expectations.

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

Consistency. Definition 1 corresponds to the definition of *weak consistency* in Jehiel (2005). The two notions do 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 have any effect on our results.

Consistency implies that, in an ABEE, the agents' expectations β_1 match the offer acceptance rate. An important feature of consistency is that histories are weighted according to the likelihood of their being reached. To see this, let $p = 0$, $n = 3$, and consider a profile σ such that every agent purchases a distribution license when he is offered the opportunity to do so in period 1 and rejects every offer made after period 1. Further, assume that under σ, in each period $t \in \{1, 2, 3\}$, every $i \in D_t$ makes an offer if he meets an agent. Thus, under σ, the first two entrants always receive an offer. The third entrant receives an offer with probability $\frac{2}{3}$ since, with probability $\frac{1}{3}$, he meets the second entrant who cannot make him an offer. 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 nodes in which agents move, including nodes in which i himself moves. This is consistent with the interpretation of i's behavior as best responding to coarse feedback on 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 can exclude the nodes 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 on the behavior in similar games in which i himself played in the past (i's actions in these past games are excluded from his analogy classes and so they influence his expectations only through their effect on r_σ). All of our results hold under both types of analogy partitions.[12]

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. It should be stressed, however, that all of our results hold when we require that the SO's strategy be optimal.

[12]Full proofs appear in an earlier version of this paper, in which the alternative analogy classes were employed (Antler, 2018).

2.2 Existence and structure of pyramid scams

The main result of this subsection (Theorem 1) shows that the SO cannot sustain a pyramid scam by means of a reward scheme that compensates the distributors only for the number of licenses they sell. Thus, behind every pyramid scam is a scheme that compensates the distributors based on at least two levels of recruitments. Theorem 2 provides necessary and sufficient conditions on the number of potential participants n and their tendency to propagate information by word of mouth p under which the SO can sustain a pyramid scam. Theorem 3 then establishes that the necessary condition of Theorem 1 is tight.

We start by asking whether the SO can sustain a pyramid scam by means of a reward scheme that compensates the distributors only for people whom they recruit to the pyramid directly. We shall refer to such a reward scheme as a one-level scheme.[13]

Definition 3 *A reward scheme R is said to be a one-level scheme if $a_\tau^R = 0$ and $b_\tau^R = 0$ for every $\tau > 1$.*

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

In a one-level scheme's induced game, each distributor's payoff depends only on the number of licenses that he sells. The proof of Theorem 1 shows that agents who understand the other agents' average behavior do not overestimate their own ability to sell licenses by much, if at all. The main challenge in the proof is to show that, in any conjectured ABEE, the last entrant who is supposed to purchase a license cannot "analogy-based" expect to sell more than one license (which, if the scheme is IC, is a necessary condition for him to purchase a license).

To get a clear intuition of the proof, suppose for a moment that $p = 0$ such that the distributors cannot access their successors' successors. Moreover, assume that the agents purchase licenses in an ABEE and that their ABEE play is symmetric. Since the likelihood of meeting new entrants decreases as time progresses, there is a period t such that each agent who receives an offer to purchase a license up to t accepts it and each agent who receives such an offer after t rejects it. Nonetheless, the distributors continue making offers after period t as they wrongly believe that the other agents might accept them. Denote the expected number of offers that each distributor makes after period t by v. Each offer that is accepted at $t' \leq t$ results in a distributor who, in expectation, makes v offers that are rejected after period t. Hence, the agents' analogy-based expectations, which are the proportion of accepted offers, cannot exceed $\frac{1}{1+v}$ and the t-th entrant cannot analogy-based expect to sell more than $\frac{v}{1+v}$ licenses.

[13] We include the sales commissions in the definition of one-level schemes for the sake of completeness.

When $p > 0$, there is an additional effect that is worth mentioning: the t-th entrant underestimates the number of offers that he will make (i.e., he believes that he will make $\tilde{v} \leq v$ offers). To grasp this effect, suppose that the t-th entrant, agent i, makes an offer to an agent j in period $t' > t$. This offer will be rejected by j as it is made after period t. If j is talkative, then, after he rejects the offer, he will refer his successors to i who will make them some additional offers. Agent i underestimates the likelihood that j will reject his offer and refer other agents to him as he thinks that j accepts offers with probability β_1 regardless of their timing. Thus, when $p > 0$ (i.e., when agents may be talkative), the t-th entrant expects to sell fewer than $\frac{v}{1+v}$ licenses.

We wish to stress that Theorem 1 does not depend on the subtleties of the network formation process (e.g., the finite number of periods or the assumption that only one agent enters the game in each period). Rather, it relies on the existence of a period from which point onward the agents do not purchase licenses. As long as the number of entrants in each period is nonincreasing as time progresses from some point in time, such a period will exist in every ABEE and the result of Theorem 1 will hold.

So far, we have shown that the SO cannot sustain a pyramid scam by means of a one-level scheme. We shall now provide necessary and sufficient conditions under which he can sustain such a scam by means of a reward scheme that compensates the distributors for their recruits and the recruits of other members of their downlines.

Theorem 2 *Fix $q = 0$. There exists a number $n^\star$ such that for every $n > n^\star$ there exists a number $p^\star(n) \in [0, 1)$ such that:*

1. *If $p \leq p^\star(n)$, then there exists an IC reward scheme R such that $\pi(R) > 0$.*
2. *If $p > p^\star(n)$, then there exists no IC reward scheme R such that $\pi(R) > 0$.*

Theorem 2 establishes that the SO can sustain a pyramid scam if both n is sufficiently large and p is sufficiently small. The talkativeness parameter p can be interpreted as a property of the product that disguises the underlying scam. A high value of p corresponds to contagious (or unique) products in the sense that people are excited to talk about them with their friends and acquaintances. Theorem 2 shows that an SO who wishes to initiate a pyramid scam will prefer to disguise it by means of a product that does not provoke such word-of-mouth advertisement.

Before we discuss the role that p plays in this result, it is useful to provide some intuition of the existence result. This intuition resembles the one behind the ABEE analysis of the finite-horizon centipede game (Jehiel, 2005), which we shall discuss in Section 5. The present setting is not stationary and this leads to nonstationary ABEE behavior, as the agents accept offers until some period t (if at all) and reject offers that arrive later. A "rational" economic agent will not join the pyramid in period t, knowing that in the later stages of the game it is no longer beneficial to purchase a license. However, the analogy-based reasoners in our model view the other agents' behavior as if it were time-invariant: each agent i wrongly believes that the others always accept offers with the average probability β_1, even when it is no longer beneficial to join the pyramid. This overoptimistic belief is what makes the agents join the pyramid in the hope of benefiting at the expense of the future entrants.

The fact that the talkativeness parameter has a negative effect on the SO's ability to sustain a pyramid scam may not be intuitive at first glance. Observe that the larger p is, the better the distributors' access to their successors' successors, which allows them to enjoy a larger potential clientele. This may suggest that p has a positive effect on the agents' willingness to participate in a pyramid scam. However, p has an additional negative effect on the agents' "feedback" (i.e., their analogy-based expectations), which may not be as transparent as the above-mentioned positive effect.

In order to grasp the intuition for the negative effect, note that, in equilibrium, the agents accept offers to join the pyramid early in the game and reject such offers in the game's later stages. This implies that the ratio of non-distributors to distributors is higher in the later stages of the game such that, compared to the early entrants, agents who enter late in the game are far more likely to encounter a non-distributor. The higher p is, the more likely each non-distributor is to be talkative and refer the late entrants whom he meets to a distributor who will make them an offer (which they will reject, as it is no longer beneficial to join the pyramid). In conclusion, a higher p implies that a larger number of offers are made late in the game and, since these offers are rejected, a higher value of p worsens the agents' analogy-based expectations to the extent that, when p is high, the SO cannot sustain a pyramid scam.

In light of the existence result of Theorem 2, it is natural to ask what is the minimal number of levels of the distributors' downlines that the compensation must be made contingent on for a pyramid scam to be sustained. Theorem 3 provides conditions under which the SO can sustain a pyramid scam by means of a two-level reward scheme (i.e., a scheme R such that $a_\tau^R = 0$ and $b_\tau^R = 0$ for every $\tau > 2$).

Theorem 3 *Fix $q = 0$. There exists a number $n^{\star\star} \geq n^\star$ such that for every $n > n^{\star\star}$ there exists a number $p^{\star\star}(n) \in [0, 1)$ such that if $p \leq p^{\star\star}(n)$, then there exists an IC two-level scheme R such that $\pi(R) > 0$.*

Unlike one-level schemes, two-level schemes require prospective participants to assess not only their own ability to recruit but also their recruits' respective ability. The agents' misspecified model of the other agents' behavior leads them to mispredict both of these variables. As we showed in Theorem 1, the agents 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 one-level scheme). In a similar manner, the agents do not overestimate their recruits' ability to sell licenses by much. While each of these two prediction errors is small, their accumulation allows the SO to overcome the incentive-compatibility constraint and sustain a pyramid scam.

2.3 A mixture of fully and boundedly rational agents

So far, we have assumed that all of the agents are either fully rational or analogy-based reasoners. Let us examine a population that consists of a mixture of the two types. We assume that $\lceil \alpha n \rceil$ of the agents are analogy-based reasoners and that $n - \lceil \alpha n \rceil$ of the agents are fully rational. As in Section 2.2, the analogy-based reasoners best respond to analogy-based expectations that are consistent with the profile of strategies played in equilibrium. The rational agents best respond to the equilibrium play using the accurate (i.e., not coarse) model of the world and, therefore, we expect them to reject offers to join a pyramid in instances where analogy-based reasoners are happy to accept them. Clearly, the impossibility results of Theorems 1 and 2.2 hold for every $\alpha > 0$. Proposition 2 shows that the positive results of Theorems 2.1 and 3 hold for every $\alpha > 0$.

Proposition 2 *Fix $q = 0$ and an arbitrary $\alpha \in (0, 1]$. There exists a number $n(\alpha)$ such that for every $n > n(\alpha)$, there exists a number $\hat{p}(n) \in [0, 1)$ such that if $p \leq \hat{p}(n)$, then there exists an IC two-level scheme R such that $\pi(R) > 0$.*

Proposition 2 establishes that the SO can sustain a pyramid scam even in the presence of fully rational agents. The intuition is that by taking n to be large it is always possible to "compensate" for the existence of the fully rational agents who reject offers in instances in which the analogy-based reasoners accept them.

When $\alpha < 1$ is sufficiently close to 1, there are cases in which rational agents purchase licenses in equilibrium. It should be noted that fully rational agents cannot be "scammed"; when they participate in a pyramid scam, they make a positive expected payoff at the expense of the analogy-based reasoners who, on average, incur losses.

It is worth noting that the presence of *fully rational agents can protect the analogy-based reasoners* from deceptive practices. The fact that they reject offers in instances in which the analogy-based reasoners will accept them worsens the analogy-based reasoners' expectations to the extent that, for small values of α (fixing an arbitrary population size $\bar{n}$), the SO cannot sustain a pyramid scam regardless of p.

2.4 Example: A multilevel scheme maximizes the SO's profit

This example illustrates that maximizing the SO's expected ABEE profit may require compensating the distributors based on strictly more than two levels of recruitments.

We shall say that an IC scheme R is *profit-maximizing* if there exists no IC scheme R' such that $\pi(R') > \pi(R)$. Note that due to their risk neutrality, both the agents and the SO can benefit from raising the stakes of the contract (i.e., multiplying the commissions and the entry fees by a constant $\gamma > 1$ can increase the SO's expected profit without changing the set of ABEEs). To bound these stakes such that a profit-maximizing scheme will exist, we shall modify the model by assuming that the maximal amount that each agent can pay for a license is $B > 0$. Under this assumption, if the SO can sustain a pyramid scam, then every profit-maximizing scheme charges a fee of B. Without loss of generality, we shall restrict our attention to such schemes.

To simplify the analysis, we set $p = 0$ such that each distributor finds it optimal to make an offer to every agent that he meets. In every ABEE that maximizes the SO's expected profit, there is some period $k < n$ such that the agents accept every offer up to period k and reject every offer made after period k. At the optimum, the SO does not make offers after period k as such offers are rejected and negatively affect the agents' analogy-based expectations. In expectation, each distributor interacts with $\sum_{t=k+1}^{n} \frac{1}{t}$ agents after period k. Thus, each accepted offer results in a distributor who, in expectation, makes $\sum_{t=k+1}^{n} \frac{1}{t}$ offers that are rejected after period k. Hence, $\beta_1 = \frac{1}{1+\sum_{j=k+1}^{n} \frac{1}{j}}$ and $\beta_2 = 1$ are consistent with the players' behavior at the optimum.

Fix a profile of strategies σ that satisfies the above properties. Note that σ pins down the agents' analogy-based expectations and the SO's expected revenue, which is the expected number of distributors multiplied by B. The profile σ can be part of an ABEE in multiple IC schemes' induced games. We shall look for the scheme that minimizes the SO's costs (given σ) in this class of schemes.

The SO's dual problem is to minimize $\kappa_1 a_1^R + \kappa_2 a_2^R + ... + \kappa_{n-k} a_{n-k}^R$ subject to the incentive-compatibility constraints that $a_1^R, ..., a_{n-k}^R \leq B$, subject to $a_1^R, ..., a_{n-k}^R \geq 0$, and subject to the k-th entrant being willing to pay $B + c$ for a license (i.e., σ being an ABEE of $\Gamma(R)$). Each weight κ_τ represents the expected cost that is associated with an increase in the commission a_τ^R given σ. The k-th entrant's willingness to pay for a license is $\sum_{\tau=1}^{n-k} \beta_1^\tau l_\tau a_\tau^R$, where l_τ is the expected number of agents at the τ-th level of the subtree of G rooted at the k-th entrant. To find an IC profit-maximizing scheme, it is sufficient to find a cost-minimizing scheme for every profile σ that satisfies the above properties and compare the SO's expected profit in the corresponding ABEEs.

For $n = 100$, $B = 1$, and $c = 0$, the first four entrants purchase a license in the ABEE in which the SO's expected profit is maximized (i.e., $k = 4$) and every profit-maximizing scheme R pays $a_1^R = 0.823$ and $a_\tau^R = 1$ for every $\tau \in \{1, ..., n-k\}$. The maximal expected profit that the SO can obtain in an ABEE is 1.922 in this case. If the SO were restricted to using a two-level scheme instead of a multilevel one, then the maximal expected profit he could obtain in an ABEE would be 1.864. Thus, the SO is strictly better off using a multilevel scheme rather than a two-level one.

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

In this section, we shall explore a "legitimate" world in which the good is intrinsically valued such that the SO can benefit from the distributors' retail sales. We shall analyze two settings that correspond to the ones that we studied in Section 2. In Section 3.1, the agents do not err, such that the SO cannot benefit from deceptive practices. In Section 3.2, the agents are boundedly rational and the SO can benefit from their mistakes as well as from their sales. In both settings, we shall focus on two controversial components of the reward schemes, namely, entry fees and recruitment commissions. We shall investigate whether or not the SO must use these components in order to maximize his expected profit.

3.1 MLM with fully rational agents

We set $q > 0$ in order to capture that the good is genuine and we use subgame perfect Nash equilibrium (SPE) to solve the model, as the agents are fully rational. As in the example in Section 2.4, we shall say that an IC reward scheme R is profit-maximizing if there exists no IC reward scheme R' such that $\pi(R') > \pi(R)$. Theorem 4 shows that if the SO produces a genuine good and faces fully rational agents, then profit-maximizing schemes need not charge entry fees or pay recruitment commissions.

Theorem 4 *Let $c > 0$. There exists an IC profit-maximizing scheme $R^\star$ such that $\phi^{R^\star} = 0$ and $a_\tau^{R^\star} = 0$ for every $\tau \geq 1$.*

The proof shows that for every IC scheme R, there exists an IC scheme $R^\star$ that does not pay recruitment commissions, does not charge entry fees, and where $\pi(R^\star) \geq \pi(R)$. Under $R^\star$, whenever a distributor interacts with an agent he is indifferent between recruiting him and not doing so. This indifference has two implications. First, in expectation, each distributor is paid as if he does not recruit at all. Each distributor can secure this payment in $\Gamma(R)$ by selling the good without recruiting as $b_1^{R^\star} \leq b_1^R$. Hence, the SO's expected transfer to each distributor is less under $R^\star$ than it is under R. The second implication of the distributors' indifference is that any profile of strategies that constitutes an SPE of $\Gamma(R)$ also constitutes an SPE of $\Gamma(R^\star)$. Thus, the scheme $R^\star$ sustains the SO's preferred SPE behavior in $\Gamma(R)$ while maintaining a lower overhead.

A scheme that charges entry fees or pays recruitment commissions cannot be profit-maximizing except in a few extreme cases (e.g., when $p = 1$, the SO's expected profit is maximized by selling directly and every IC scheme is profit-maximizing). In fact, if an SPE of $\Gamma(R)$ results in two or more distributors and R charges entry fees or pays recruitment commissions, then $b_1^{R^\star} < b_1^R$ and $\pi(R^\star) > \pi(R)$. Thus, in any case where the SO uses at least two distributors, schemes that charge entry fees or pay recruitment commissions are not profit-maximizing

The compensation under $R^\star$ is based on *more than two levels of retail sales*. The commission $b_1^{R^\star}$ is meant to cover the cost c. The purpose of $b_2^{R^\star}$ is to incentivize information propagation. This incentive is necessary since a distributor who recruits an agent is basically recruiting a competitor, as he loses access to that agent's successors. However, $b_2^{R^\star}$ may not suffice to incentivize recruitment of new competitors since it does not compensate the distributors for the fact that these competitors may recruit additional competitors themselves. For example, a distributor d who recruits an agent i loses access to i's successors. As long as i does not recruit anyone, $b_2^{R^\star}$ compensates d for these losses. However, if i recruits an additional agent j, then the recruitment reduces i's sales as he can no longer sell to j's successors and this may negatively affect d's reward, which, in a two-level scheme, is based on i's sales but not on j's sales. The higher-order commissions (e.g., $b_3^{R^\star}$) compensate the distributors for the fact that their recruits may recruit new recruits (who may in turn recruit, and so on).

So far in this section, we have shown that recruitment commissions and entry fees are inconsistent with profit maximization when the good is intrinsically valued and the agents are fully rational. We now complete the analysis by exploring a setting in which the SO has two potential sources of profit: selling the good and scamming the agents.

3.2 MLM with genuine goods and analogy-based reasoners

In this section, the agents are boundedly rational such that the SO can exploit them. Unlike in Section 2.2, we shall assume that the agents are willing to pay for the good (i.e., $q > 0$). As in Section 3.1, we shall examine whether or not profit-maximizing schemes pay recruitment commissions or charge entry fees.

We shall simplify the analysis by assuming that $p = 0$ such that the distributors do not have access to their successors' successors. To guarantee that profit-maximizing schemes exist, we shall assume, as in the example in Section 2.4, that each agent cannot pay more than $B > 0$ for a license. Unlike in the example, profit-maximizing schemes do not necessarily charge a fee of B.

Proposition 3 shows that even though the agents are vulnerable to deceptive practices, when the potential gains from sales are large, schemes that pay recruitment commissions or charge entry fees are not profit-maximizing.

Proposition 3 *Fix $q > 0$, $c > 0$, and $p = 0$. There exists a number $\hat{n}$ such that if $n > \hat{n}$ and R is a profit-maximizing IC scheme, then $\phi^R = 0$ and $a_\tau^R = 0$ for every $\tau \geq 1$.*

When n is large, the potential gains from trade are large as well, which makes the distributors more valuable to the SO. The first part of the proof confirms this intuition. It shows that for any pair, $\epsilon > 0$ and $\rho > 0$, there is a number $n(\epsilon, \rho)$ such that, for every $n > n(\epsilon, \rho)$, every ABEE that maximizes the SO's expected profit induces a number of distributors greater than ρ with probability greater than $1 - \epsilon$.

The second part of the proof examines the SO's "dual" problem: finding an IC scheme that minimizes the SO's expected cost given an entry fee ϕ, a price $\eta = 1$, and a profile σ, and that is subject to the requirement that σ be part of an ABEE of the scheme's induced game. Denote a scheme that solves this problem for ϕ and σ by $R(\phi, \sigma)$. We show that $\pi(R(0, \sigma)) > \pi(R(\phi, \sigma))$ for every pair (ϕ, σ) such that $\phi > 0$ and the number of agents who purchase a license under σ is sufficiently large. Thus, charging entry fees is inconsistent with profit maximization when many agents purchase a license. Hence, for large values of n, IC schemes that charge entry fees (or pay recruitment commissions, which require charging entry fees) are not profit-maximizing.

Let us consider the main effects of charging a fee on the SO's expected profit. First, the fee raises the SO's revenue per distributor. Second, it raises the cost of becoming a distributor $\phi + c$ such that maintaining the agents' willingness to purchase a license requires raising the commissions and the SO's costs. When the number of distributors is large, the second effect dominates the first effect as the SO has to pay multiple commissions for the bulk of the recruitments and sales (such that raising the commissions to compensate the distributors for the higher fee becomes very costly). Hence, if the number of distributors is likely to be large, the SO's expected profit is maximized by means of a reward scheme that does not charge entry fees.

4 Comparison to Prominent Behavioral Models

The agents in our model correctly predict the number of individuals that they will encounter and have a correct, yet coarse, estimate of the likelihood that these individuals will accept offers. One important motivation for these assumptions is that information about the time the scheme started, the number of current members, and coarse summary statistics on the sales force's performance in past schemes is typically easier to obtain compared to more accurate details about the likelihood of recruiting new participants in different stages of the scheme.

Inspired by the behavioral finance literature and Shiller's (2015) description of speculative bubbles as "naturally occurring Ponzi schemes," we explore the implications of prominent behavioral finance models for our results.

Non-common priors

A natural modeling assumption in the context of pyramid schemes is that people underestimate the finiteness of the process and, thus, overestimate the number of individuals that they will encounter. Let us examine a case in which different agents hold different prior beliefs about the length of the game. To do so, we relax the assumption that the number of periods is fixed and assume instead that each agent i believes that, in each period $t \in \mathbb{N}$, conditional on the game reaching period t, the game continues for an additional period (and a new agent enters) with a probability of δ_i, and that it terminates at t with a probability of $1 - \delta_i$.

We assume that the agents' priors are drawn from a distribution with a support $\left[\underline{\delta}, \bar{\delta}\right]$ and interpret this distribution as the distribution of opinions in the general population. For the sake of brevity, we set[14] $\bar{\delta} < 1$. The next result shows that the SO cannot sustain a pyramid scam unless there are additional departures from the classic rational expectations model.

Proposition 4 *Let $q = 0$. There exists no IC reward scheme R such that the SO makes a strictly positive expected profit in a subgame perfect Nash equilibrium of* $\Gamma(R)$.

Agents who hold non-common prior beliefs differ in how they evaluate the future, but their evaluation is independent of the history. However, in an ABEE, the agents' evaluation of the future is affected by the overall behavior of the other agents and, in

[14] The assumption that $\bar{\delta} < 1$ can be replaced by mild assumptions on the distribution from which the priors are drawn and the agents' knowledge. For instance, the result of Proposition 4 would hold under the assumption that it is commonly known that the priors are drawn from the uniform distribution on $[0, 1]$.

particular, by the events that take place early in the game. This history-dependence prevents the use of backward induction arguments and enables pyramid scams to work.

Extrapolation and the feedback loop theory of bubbles

A strand of the behavioral finance literature (starting with DeLong et al., 1990) suggests that behavioral biases such as the representativeness heuristic lead individuals to extrapolate from past trends when they form their expectations. Shiller (2015) coined the term "feedback loop" in reference to situations in which "past price increases generate expectations of further price increases." In the context of pyramid scams, prospective participants who extrapolate from the success of individuals who joined the pyramid in its early stages may overestimate their own ability to recruit new members to the pyramid. We now incorporate this type of behavior into our model and examine the robustness of the paper's main result.

For every $t > 1$, denote by γ_t the average number of recruitments per distributor up to period t, and set $\gamma_1 = 0$. Suppose that some of the agents are *extrapolators*. An extrapolator who receives an offer to join the pyramid in period $t \geq 1$ believes that he will recruit γ_t new agents to the pyramid. Since the SO recruits some of the agents, $\gamma_t < 1$ for every $t > 1$. Recall that a necessary condition for an agent to purchase a license in an IC one-level scheme's induced game is that he expects to sell at least one license. It follows that the SO cannot sustain a pyramid scam by means of an IC one-level scheme even in the presence of extrapolating agents.[15]

5 Concluding Remarks

Legitimate multilevel marketing 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 allows us to draw the boundary between the two based on observable properties of their underlying compensation structures. We provided necessary and sufficient conditions under which it is possible to sustain a pyramid scam even though the potential participants' beliefs are statistically correct. The main result of the paper makes a connection between dubious "passive income" promises and pyramid scams by showing that every pyramid scam has a reward scheme that compensates the participants for at least two levels of recruitments. That is, schemes that compensate the participants only for people whom they recruit to the sales force cannot sustain a pyramid scam, but multilevel schemes can sustain such a scam.

[15] Whether extrapolating agents purchase a license in a game that is induced by a multilevel scheme depends on their expectations about their successors' ability to recruit new members to the pyramid.

We showed that even in the presence of fully rational agents, the SO may sustain a pyramid scam. However, their presence makes it more difficult for the SO to sustain such a scam. They reject offers to join a pyramid in instances where boundedly rational agents would be happy to join and, therefore, their presence worsens the boundedly rational agents' analogy-based expectations to the extent that it may be impossible for the SO to sustain the pyramid scam. Thus, the presence of fully rational agents can prevent the boundedly rational agents from participating in a pyramid scam.

Pyramid scams rely on two components, namely, recruitment commissions and entry fees. We established that legitimate companies (i.e., ones that promote genuine goods) that face fully rational consumers cannot benefit from basing their compensation on these pyramidal components. In such instances, it is possible to maximize the company's expected profit while compensating the sales force only for sales. In fact, even when such a company faces boundedly rational agents who are vulnerable to deceptive practices, it may find it suboptimal to use these pyramidal components if the demand for its product is sufficiently large.

It is well known that in the unique SPE of the finite-horizon centipede game, players always *stop* even though they can benefit if they *continue* for a few more rounds. Jehiel (2005) resolves this paradox by showing that it is possible to sustain an ABEE in which players continue. If the game is sufficiently long, then under the coarsest partition there exists an ABEE in which the players continue until the last round. Such an ABEE exists even if only one of the players has a coarse reasoning as the players generate surplus along the path of play. There are several other differences between the games that we analyzed in this article and the centipede game analyzed by Jehiel that not only allow us to answer questions related to the application, but also lead us to results that are fundamentally different, as is illustrated in our impossibility results.

The effect of p on the agents' analogy-based expectations, and the restrictions on the payoff function that result both from incentive compatibility and from the requirement that the SO's expected profit be positive, may prevent the SO from sustaining a pyramid scam even if the game is arbitrarily long (e.g., if $p = 1$). Moreover, they impose an interesting structure on the participants' compensation when a scam is sustainable.

We shall conclude by discussing a few extensions and modifications of the model.

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 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 R such that $\sum_{\tau=1}^{n} a_{\tau}^{R} \leq \phi^{R}$ and $\sum_{\tau=1}^{n} b_{\tau}^{R} \leq \eta^{R}$ 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 illustrate that, even under the stronger incentive constraint, one-level schemes cannot sustain a pyramid scam whereas two-level schemes can sustain such a scam.

Extension: A limited number of recruitment opportunities

Individuals who join a pyramid may find it natural to first approach their immediate friends as approaching strangers is perhaps more difficult. Such individuals exhaust their best opportunities to recruit new members to the pyramid soon after they join. In order to roughly approximate this, we modify the network formation model such that each player can meet new agents only in the first $x > 0$ periods after he enters the game (i.e., the time-t entrant can meet new agents in periods $t+1, ..., t+x$). Moreover, we assume that in each period $t \in \{1, ..., n\}$, $\mu_t > 0$ new agents enter the game. As in the baseline model, the player who meets the entrant is drawn by nature uniformly at random. Observe that in the general network formation model, agents do not necessarily meet fewer agents than their successors.

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

Set $q = p = 0$ and $x = 1$ and consider a profile in which every agent who receives an offer up to period $k < n$ accepts it and every agent who receives an offer after period k rejects it. Moreover, assume that every distributor who meets an agent makes him an offer. Since $p = 0$, no offer is made after period $k+1$. Observe that every symmetric ABEE in which the SO's expected profit is strictly positive must have this structure when $p = 0$ and $x = 1$. Under this profile, $\mu_1 + ... + \mu_{k+1}$ offers are made and $\mu_1 + ... + \mu_k$ of them are accepted. Hence, $\beta_1 = \frac{\mu_1 + ... + \mu_k}{\mu_1 + ... + \mu_{k+1}}$ is 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 $z-1$. Agent i believes that, in expectation, he will 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 one-level scheme, will not cover the cost of becoming a distributor.

The profile that we described above can be part of an ABEE of a two-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 $a_1^R = a_2^R = 0.5\phi^R$, an agent who enters the game in period 1 analogy-based expects a reward of $1.2\phi^R - \phi^R - c$. An agent who enters the game in period 2 analogy-based expects to sell $\beta_1 \frac{\mu_3}{\mu_2} = 2$ licenses and, hence, he analogy-based expects a reward of $-c$ such that he finds it optimal not to purchase a license. Thus, for large ϕ^R, we have an IC two-level scheme R (that satisfies $a_1^R + a_2^R \leq \phi^R$) such that $\pi(R) > 0$.

Uncertainty about arrival time

Abreu and Brunnermeier (2003) study a model in which a finite process creates a bubble that bursts after a synchronized attack by a sufficient number of investors or at the end of the process. The investors in their model become aware of the bubble sequentially, and face *uncertainty about the time at which the bubble started* and how many other investors are aware of the bubble. They show that the bubble may persist long after all of the investors are aware of its existence.

Uncertainty about arrival time *cannot lead to participation* in a pyramid scam when there are no deviations from the classic rational expectations model. The reason that such uncertainty sustains a bubble in Abreu and Brunnermeier's model is that, unlike a pyramid scam, their trading game is not a zero-sum game. Underlying the bubble in their model is an exogenous process that represents behavioral traders. Rational investors who ride the bubble profit at the expense of those behavioral traders, unlike pyramid-scam participants who, on average, incur losses.

What would be the implications of assuming that agents are uncertain about the time at which they enter the game, in addition to assuming that their reasoning is coarse? Adding a small noise to the model would not change the essence of our main results. However, if the agents' information about their time of entry were to become very coarse, then it would be impossible for the SO to sustain a pyramid scam. The reason for this, loosely speaking, is that when an agent is completely ignorant about his position in the game tree, best responding to the other agents' average behavior (as in an ABEE) is a relatively small mistake.

Gradual arrival

The gradualness in the arrival of agents plays a key role in the model as it affects the number of opportunities the agents have to recruit. If the agents arrive "too fast" then there is not enough variation in the agents' expected number of successors and so the SO cannot sustain a pyramid scam. For example, in the extreme case in which all of the agents enter the game in the same period, none of them could recruit and, therefore, none of the agents would be willing to participate in the pyramid.

Multiple schemes

In the model, there is one scheme by which the agents can become distributors. In practice, MLM companies often offer several such schemes. Can the SO benefit from offering multiple schemes and having the agents self-select into them? Potentially, this would enable the SO to code the nonstationarity of the setting into the rewards.

When an SO who offers a menu of schemes faces a boundedly rational agent, they can disagree on which scheme is more profitable for the agent (i.e., which scheme induces greater transfers from the SO). Thus, there are cases where, in each period, a boundedly rational agent mistakenly chooses the scheme that is most profitable for the SO and least profitable for the agent, such that offering multiple schemes renders the SO better off. However, if the SO faces a fully rational agent, they always agree on which scheme is more profitable for the agent such that offering multiple schemes can only render the SO worse off compared to offering one profit-maximizing scheme.[16]

In conclusion, when the SO faces fully rational agents he cannot increase his profit by means of offering multiple schemes and having the agents self-select into these schemes; however, the SO can benefit from offering multiple schemes for joining the sales force when he faces agents who are vulnerable to deceptive practices as they may self-select into the "wrong" schemes.

References

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

[2] Antler, Y. (2018): "Multilevel Marketing: Pyramid-Shaped Schemes or Exploitative Scams," CEPR Discussion Paper No. DP13054.

[3] Athreya, K. B. and Ney, P. E. (1972): *Branching Processes,* New York, Springer.

[4] 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.

[16]The proof is similar to the proof of Theorem 4, where R is replaced by a menu of IC schemes $\mathcal{S}$ and b_1^R is replaced by $max\left\{b_1^{R'} : R' \in \mathcal{S}\right\}$.

[5] 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.

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

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

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

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

[10] 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.

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

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

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

[14] Eliaz, K. and Spiegler, R. (2007): "A Mechanism-Design Approach to Speculative Trade," *Econometrica,* 75, 875–884.

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

[16] Eliaz, K. and Spiegler, R. (2008): "Optimal Speculative Trade among Large Traders," *Review of Economic Design,* 12(1), 45–74.

[17] Eliaz, K. and Spiegler, R. (2009): "Bargaining over Bets," *Games and Economic Behavior*, 66, 78–97.

[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] Eyster, E. and Rabin, M. (2010): "Naïve Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2, 221–243.

[22] *FTC v. Fortune Hi-Tech Marketing, Inc.* 13 CV 123 KSF (N.D. Ill, 2014).

[23] *FTC v. BurnLounge, Inc.* 753 F.3d 878 (9th Cir., 2014).

[24] 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/news-events/press-releases/2016/07/herbalife-will-restructure-its-multi-level-marketing-operations.

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

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

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

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

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

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

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

[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, Caltech.

[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 09.

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

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

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

[45] 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.

[46] Parloff, R. (2015): "The Siege of Herbalife," *Fortune*, September 09.

[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] 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.

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

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

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

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

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

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

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

Appendix A: Proofs

Proof of Proposition 1. Assume by way of contradiction that there exists an IC reward scheme R such that $\pi(R) > 0$. Since $q = 0$, $\pi(R) > 0$ implies that $\phi^R > 0$ and that there exists a subgame perfect Nash equilibrium (SPE) of $\Gamma(R)$ in which at least one of the agents purchases a license. Denote by k the last period in which an agent purchases a license in any SPE of $\Gamma(R)$. Since R is IC, $\eta^R = 0$ implies that $b_1^R = b_2^R = ... = 0$. As $q = 0$, the agents do not purchase the good if $\eta^R > 0$. This leads to a contradiction as, in every SPE, the k-th entrant obtains a payoff of $-c - \phi^R < 0$ if he purchases a license and, hence, must refrain from doing so.

Proof of Theorem 1. The proofs of Theorems 1 and 2 rely on the following lemma.

Lemma 1 *Let (σ, β) be an ABEE in which every agent rejects every offer that he receives after period $k^\star$, where $k^\star \leq n-1$. In (σ, β), every distributor who interacts with an agent after period $k^\star$ makes him an offer.*

Proof. Consider (σ, β). Since $\beta_1 a_1^R \geq 0$, every distributor who interacts with an agent in period n makes him an offer. Consider $t \in \{k^\star, ..., n-1\}$ and assume that, in every period $t' > t$, every distributor who interacts with an agent makes him an offer. To prove the lemma by induction, we need to show that every distributor who interacts with an agent in period t makes him an offer.

Consider an agent i who holds a distribution license at the end of period t. In expectation, i will meet $\sum_{j=t+1}^{n} \frac{1}{j}$ agents after period t. Denote by S the set of these agents and their successors in G. By the induction hypothesis, i makes an offer to every $j \in S$ with whom he interacts. Since $t \geq k^\star$, every offer that i makes after period t is rejected. Thus, i interacts with $j \in S$ if and only if $\psi_l = 1$ for every agent l on the directed path connecting i and j, an event that occurs with probability $p^{d_G(i,j)-1}$. Hence, the expected number of offers that i makes after period t is at least[17] v_t, where

$$v_t = \sum_{j=t+1}^{n} \frac{1}{j} + p \sum_{j=t+1}^{n-1} \sum_{j'=j+1}^{n} \frac{1}{jj'} + p^2 \sum_{j=t+1}^{n-2} \sum_{j'=j+1}^{n-1} \sum_{j''=j'+1}^{n} \frac{1}{jj'j''} + ... \qquad (1)$$
$$+ ... + p^{n-t-1} \frac{1}{(t+1) \times (t+2) \times ... \times (n-1) \times n}$$

Note that every accepted offer results in a distributor and that no offer is accepted after period t. Thus, for every accepted offer, there are, in expectation, at least v_t rejected offers. Hence, the proportion of accepted offers, β_1, cannot exceed $\frac{1}{1+v_t}$.

A distributor j is said to be a member of the λ-th level of distributor i's *downline* if $j \in G_i$ and there are $\lambda - 1$ distributors on the directed path connecting i to j. Denote by $l_{\lambda,t}$ the *analogy-based expected* (i.e., assuming that after period t every distributor makes offers with probability β_2 and each agent accepts offers with probability β_1) number of distributors in the λ-th level of the downline of an agent who purchases a license in period t. Set $l_{0,t} := 1$ and note that $l_{\lambda,t}$ is weakly decreasing in t.

Suppose that a distributor i interacts with an agent j in period t. If i does not make an offer to agent j (or if j rejects i's offer) and $\psi_j = 1$, then i will be able to make offers to j's successors. By the induction assumption, i will make an offer to each of j's successors with whom he interacts. Note that i believes that each of these offers will be accepted with probability $\beta_1 \geq 0$. Thus, i "analogy-based" expects to make $p\hat{v}_t$ offers to j's successors if he does not sell a license to j, where

$$\hat{v}_t = \sum_{j=t+1}^{n} \frac{1}{j} + (1-\beta_1)\, p \sum_{j=t+1}^{n-1} \sum_{j'=j+1}^{n} \frac{1}{jj'} + (1-\beta_1)^2 p^2 \sum_{j=t+1}^{n-2} \sum_{j'=j+1}^{n-1} \sum_{j''=j'+1}^{n} \frac{1}{jj'j''} + ... \qquad (2)$$
$$... + (1-\beta_1)^{n-t-1} p^{n-t-1} \frac{1}{(t+1) \times (t+2) \times ... \times (n-1) \times n} \leq v_t$$

Hence, i analogy-based expects to sell $p\hat{v}_t\beta_1 \leq \frac{v_t}{1+v_t} < 1$ licenses to j's successors in case he does not sell a license to j. Each sale to j's successors increases the analogy-based expected number of distributors in the λ-th level of i's downline by less than $l_{\lambda-1,t}$ since $l_{\lambda-1,t}$ is weakly decreasing in t. A sale to j increases the analogy-based expected number of distributors in the λ-th level of i's downline by $l_{\lambda-1,t}$. Hence, selling a license

[17] If i entered the game before period t, then, in addition to the offers that he makes to the members of S after period t, agent i can make offers to agents who join $G_i - S$ after period t.

to j results in a weakly greater expected number of distributors in every level of i's downline compared to not selling a license to j. It follows that selling a license to j results in a weakly greater expected reward compared to not selling a license to j. ■

Assume by way of contradiction that there exists an IC one-level scheme R such that $\pi(R) > 0$. Since $\pi(R) > 0$ and $q = 0$, it follows that $\phi^R > 0$ and that there exists an ABEE (σ, β) of $\Gamma(R)$ in which the agents purchase licenses. Denote by t the last period in which an agent purchases a license in (σ, β). By Lemma 1, every agent who purchases a license in period t makes an offer to every agent with whom he interacts. In expectation, such an agent makes v_t offers, where v_t is given in (1). As we showed in Lemma 1, $\beta_1 \leq \frac{1}{1+v_t}$. The t-th entrant "analogy-based" expects to make $\hat{v}_t$ offers and to sell $\hat{v}_t \beta_1$ licenses if he purchases a license, where $\hat{v}_t$ is given in (2). Since R is an IC one-level scheme and $q = 0$, the t-th entrant analogy-based expects a payoff of at most $\hat{v}_t \beta_1 a_1^R - \phi^R \leq v_t \frac{1}{1+v_t} \phi^R - \phi^R < 0$ in case he becomes a distributor. This is in contradiction to the optimality of purchasing a license in period t.

Proof of Theorem 2. The proof consists of three parts. First, we show that if $p = 1$, then the SO cannot sustain a pyramid scam regardless of the value of n. The second part of the proof shows that if the SO can sustain a pyramid scam given n and $p^\star > 0$, then he can do so given n and any $p \leq p^\star$. The third part of the proof shows that there exists $n^\star$ such that if $n > n^\star$ and $p = 0$, then the SO can sustain a pyramid scam.

Part 1. Let $p = 1$ and assume by way of contradiction that there exists an IC scheme R such that $\pi(R) > 0$. Since $\pi(R) > 0$ and $q = 0$, it follows that $\phi^R > 0$ and that an ABEE (σ, β) of $\Gamma(R)$ in which the agents purchase licenses exists. Denote by t the last period in which an agent purchases a license on the equilibrium path of (σ, β). Consider an agent i who buys a license in period t. By Lemma 1, i makes an offer to every agent with whom he interacts. In expectation, i makes v_t offers, where v_t is given in (1). As we showed in the proof of Lemma 1, $\beta_1 \leq \frac{1}{1+v_t}$. Plugging $p = 1$ into (1), $v_t = \frac{n-t}{t+1}$. Since R is IC and $q = 0$, agent i's payoff cannot exceed $s\phi^R - \phi^R$, where s is the number of i's successors who purchase a license. Bhattacharya and Gastwirth (1984, p. 531) show that the expected number of successors of the t-th entrant is $\frac{n-t}{t+1}$. This contradicts the optimality of purchasing a license in period t as $\beta_1 \frac{n-t}{t+1} < 1$.

Part 2. Fix $n^\star$ and $p^\star > 0$, and suppose that there exists an IC scheme R such that $\pi(R) > 0$. Since $\pi(R) > 0$ and $q = 0$, there exists an ABEE (σ, β) of $\Gamma(R)$ in which the agents purchase licenses. We shall describe a profile σ' and an IC scheme R' such that the fact that (σ, β) is an ABEE of $\Gamma(R)$ implies that σ' is part of an ABEE of $\Gamma(R')$ that induces a strictly positive expected profit for the SO when $p = p^\star$. Subsequently, we shall show that σ' is part of such an ABEE of $\Gamma(R')$ for every $p < p^\star$.

Denote by k the last period in which an agent purchases a license in (σ, β). By Lemma 1, under σ, every distributor makes an offer to every agent with whom he interacts after period k. In expectation, an agent who purchases a license in period k makes $v_k(p^\star)$ offers, where $v_k(p^\star)$ is given in (1) for $t = k$ and $p = p^\star$ (we write $v_k(p)$ explicitly as a function of p as we shall vary the value of p later on). As we showed in the proof of Lemma 1, consistency implies that $\beta_1 \leq \frac{1}{1+v_k(p^\star)}$.

Consider a profile of strategies σ' in which: (i) the SO makes an offer to every agent with whom he interacts in period k and never makes offers in other periods, (ii) every distributor makes an offer to every agent with whom he interacts, and (iii) every agent accepts every offer that he receives up to period k and rejects every offer that he receives after period k. Denote the analogy-based expectations that are consistent with σ' when $p = p^\star$ by β_1' and β_2'. Clearly, $\beta_2' = 1$.

Consider an arbitrary agent i who purchases a license in period k. According to σ', i makes an offer to every agent with whom he interacts. In expectation, he makes $v_k(p^\star)$ offers after period k, where $v_k(p^\star)$ is given in (1) for $t = k$ and $p = p^\star$. Thus, under σ', every offer that is made in period k is accepted and results in a distributor who, in expectation, makes $v_k(p^\star)$ offers. Since no offer is made prior to period k and the SO does not make offers after period k, it follows that $\beta_1' = \frac{1}{1+v_k(p^\star)} \geq \beta_1$.

Let R' be a scheme such that $a_\tau^{R'} = x\phi^{R'}$ and $b_\tau^{R'} = 0$ for every $\tau \geq 1$, $x \leq 1$, and $\phi^{R'} = \phi^R$. Since $x \leq 1$, R' is IC. Given σ', the SO's expected profit in $\Gamma(R')$ is at least $\frac{\phi^R}{k}$. We now show that there exists $x \leq 1$ such that $(\sigma', \beta_1', \beta_2')$ is an ABEE of $\Gamma(R')$.

Consider the distributors' behavior. For every $x \leq 1$, each distributor i's objective in $\Gamma(R')$ is to maximize the expected number of his successors who purchase a license. Since $\beta_2' = 1$, making an offer to every agent with whom i interacts maximizes this number according to i's analogy-based expectations.

Consider the decision whether or not to purchase a license. It suffices to find $x \leq 1$ such that the k-th entrant is indifferent whether to purchase a license or not as he expects a greater (respectively, smaller) payoff than any agent who buys a license after (respectively, before) period k. Since (σ, β) is an ABEE of $\Gamma(R)$, the k-th entrant finds it optimal to purchase a license given R, β_1, and β_2. Since R is IC, for a sufficiently large $x \leq 1$, the k-th entrant finds it optimal to purchase a license given R', β_1, and β_2. Given R', increasing β_1 and β_2 to β_1' and β_2' only raises the k-th entrant's payoff according to his analogy-based expectations. Thus, we can adjust $x \leq 1$ such that σ' is part of an ABEE of $\Gamma(R')$.

Set $n^\star$ and an arbitrary $p < p^\star$. To complete the proof, we shall show that there exists $x \leq 1$ such that σ' is part of an ABEE of $\Gamma(R')$ for $p < p^\star$. Note that the

the optimality of the distributors' behavior under σ' and the lower bound for the SO's expected profit do not depend on the value of p. Hence, all we need to show is that there exists $x \leq 1$ such that the k-th entrant finds it optimal to purchase a license given p (it will then be easy to adjust x such that the later entrants do not purchase a license).

The distributors' rewards in $\Gamma(R')$ are linear in the number of their successors who purchase a license. Recall that the k-th entrant finds it optimal to purchase a license given $x \leq 1$ and $p^\star$ (i.e., σ' in an ABEE of $\Gamma(R')$ for $p = p^\star$). To complete the proof, we shall show that the expected number of the k-th entrant's successors who purchase a license according to his analogy-based expectations that are induced by σ' and p is weakly decreasing in p.

Let $l_z = E[|\{j \in G_{i_k} : d_G(i_k, j) = z\}|]$ be the expected number of agents in the z-th level of the subtree of G rooted at the k-th entrant. For example, $l_1 = \sum_{j=k+1}^{n} \frac{1}{j}$ and $l_2 = \sum_{j_1=k+1}^{n-1} \sum_{j_2=j_1+1}^{n} \frac{1}{j_1 j_2}$. Note that $v_k(p) = l_1 + pl_2 + p^2 l_3 + ... + p^{n-k-1} l_{n-k}$ and that $\beta_1'(p) = \frac{1}{1+v_k(p)}$ is consistent with σ' and p. The expected number of the k-th entrant's successors who purchase a license given p, σ', $\beta_1'(p) = \frac{1}{1+v_k(p)}$ and $\beta_2' = 1$ is:

$$\frac{l_1}{(1+v_k(p))} + \frac{l_2(1+pv_k(p))}{(1+v_k(p))^2} + \frac{l_3(1+pv_k(p))^2}{(1+v_k(p))^3} + \frac{l_4(1+pv_k(p))^3}{(1+v_k(p))^4} + ... \qquad (3)$$

The proof that the derivative of (3) w.r.t. p is negative appears in the Technical Results section in Appendix B.

Part 3. This part of the proof shows that there exists $n^\star$ such that if $n > n^\star$ and $p = 0$, then there exists an IC scheme R such that $\pi(R) > 0$. This part follows directly from a more general result given in Proposition 2 and, therefore, its proof is omitted.

Proof of Theorem 3. Theorem 3 follows as a corollary of Proposition 2 for $\alpha = 1$.

Proof of Proposition 2. Set $p = 0$ and choose an arbitrary $\alpha \in (0, 1]$. Consider a profile of strategies σ such that every analogy-based reasoner accepts (respectively, every rational agent rejects) every offer that he receives in period $t = 1$, every agent rejects every offer that he receives in each period $t > 1$, the SO makes an offer only in period $t = 1$, and every distributor makes an offer to every agent with whom he interacts. Under σ, the SO's expected profit is $\frac{\lceil \alpha n \rceil}{n} \phi^R$. We shall find an IC scheme R such that $\phi^R > 0$ and σ is part of an ABEE of $\Gamma(R)$.

Since every distributor makes an offer to every agent with whom he interacts, $\beta_2 = 1$ is consistent with σ. As the SO's period-1 offer is accepted by analogy-based reasoners and rejected by fully rational agents, in expectation, $\frac{\lceil \alpha n \rceil}{n}$ offers are accepted in period 1. Each of them results in a distributor. Plugging $t = 1$ and $p = 0$ into (1), each distributor makes, in expectation, $\sum_{t=2}^{n}\frac{1}{t}$ offers after period 1. Since all of these offers are rejected, $\beta_1 = \frac{\frac{\lceil \alpha n\rceil}{n}}{1+\frac{\lceil \alpha n\rceil}{n}\sum_{t=2}^{n}\frac{1}{t}}$ is consistent with σ.

Consider an IC two-level reward scheme R such that $\phi^R > 0, a_1^R = a_2^R = x\phi^R, x \leq 1$, and $b_1^R = b_2^R = 0$. Since $p = 0$, under σ, every analogy-based reasoner i believes that, as a distributor, he will make an offer to every agent $j \in G_i$ such that $d_G(i,j) = 1$ and only to these agents. Since $p = 0$ and $\beta_2 = 1$, agent i also believes that each agent $j \neq i$ who purchases a license will make an offer to every agent $l \in G_j$ such that $d_G(j,l) = 1$ and only to these agents. Hence, an analogy-based reasoner who purchases a license in period 1 believes that, in expectation, he will sell $\beta_1 \sum_{t=2}^{n}\frac{1}{t}$ licenses and that the agents who buy these licenses will sell, in expectation, $\beta_1^2 \sum_{t=2}^{n-1}\sum_{t'=t+1}^{n}\frac{1}{tt'}$ licenses.

As the harmonic sum diverges, $\lim_{n\to\infty}\beta_1\sum_{t=2}^{n}\frac{1}{t} = 1$ and $\lim_{n\to\infty}\beta_1^2\sum_{t=2}^{n-1}\sum_{t'=t+1}^{n}\frac{1}{tt'}$ $= \frac{1}{2}$. Thus, for a sufficiently large n, an analogy-based reasoner who purchases a license in period 1 expects a payoff arbitrarily close to $\frac{3x\phi^R}{2} - \phi^R - c$. Since the likelihood of meeting the new entrant decreases as time progresses, he expects a smaller payoff if he purchases a license in period $t > 1$. It is possible to choose x and ϕ^R such that it is optimal for every analogy-based reasoner to purchase a license at $t = 1$ and it is not optimal to do so at any $t > 1$. To complete the proof that σ is part of an ABEE of $\Gamma(R)$ note that, since $p = 0$, every distributor who meets an agent finds it optimal to make him an offer (regardless of x). Moreover, fully rational agents find it optimal to reject every offer as such agents correctly believe that they will not sell licenses. Thus, there exists an IC two-level scheme R such that $\pi(R) > 0$.

Proof of Theorem 4. Consider an arbitrary IC scheme R such that[18] $\eta^R \in (0,1]$. Step 1 shows that there exists an IC scheme $R^\star$ such that $\phi^{R^\star} = 0$, $a_\tau^{R^\star} = 0$ for every $\tau \geq 1$, and $\pi(R^\star) \geq \pi(R)$. Step 2 shows that a profit-maximizing scheme exists.

Step 1. Let σ be an SPE of $\Gamma(R)$ that induces an expected profit of $\pi(R)$ for the SO. If no agent purchases a license on the equilibrium path of σ, then $\pi(R') \geq \pi(R)$ for every scheme R' as the SO can always refrain from selling licenses. Suppose that there exists an agent who purchases a license on the equilibrium path of σ and let k be the last period in which an agent purchases a license on that path.

For every $t \geq 1$, if the t-th entrant purchases a license but does not sell licenses, then, in expectation, he sells qv_t units of the good, where v_t is given in (1). Since the k-th entrant does not sell licenses in σ, it follows that $qv_k b_1^R \geq c + \phi^R$. Note that $v_t > v_k$ for every $t < k$ and that the optimality of the t-th entrant's strategy implies that he obtains an expected payoff of at least $qv_t b_1^R - \phi^R - c$ if he purchases a license.

[18]If $\eta^R \notin (0,1]$, then no agent ever purchases a license in an SPE of $\Gamma(R)$. Since the SO can always refrain from selling licenses, $\pi(R) \leq \pi(R')$ for every IC scheme R'.

Consider a scheme $R^\star$ such that $\eta^{R^\star} = \eta^R$, $qv_k b_1^{R^\star} = c$, $\phi^{R^\star} = 0$, and for every $\tau \geq 1$, it holds that $a_\tau^{R^\star} = 0$ and $b_\tau^{R^\star} = b_1^{R^\star} p^{\tau-1}$. Since $qv_k b_1^{R^\star} = c \leq c + \phi^R \leq qv_k b_1^R$, it follows that $b_1^{R^\star} \leq b_1^R$. As R is IC and $b_1^R \geq b_1^{R^\star}$, it must be that $R^\star$ is IC.

Under $R^\star$, each distributor j obtains an expected transfer of $qp^{x-y-1} b_{y+1}^{R^\star} = qp^{x-1} b_1^{R^\star}$ for every potential sale to an agent i such that $d_G(j,i) = x$ and there are y distributors on the path connecting j and i. Since this expression is independent of y, each distributor j who interacts with an agent is indifferent between selling a license to him and not doing so regardless of j's beliefs about future play. Hence, in every SPE of $\Gamma(R^\star)$, if the t-th entrant purchases a license, then, in expectation, he will obtain commissions of $qv_t b_1^{R^\star}$ as if he were not selling licenses.

Since $qv_k b_1^{R^\star} = c \leq qv_k b_1^R - \phi^R$, it holds that $qv_t b_1^{R^\star} \leq qv_t b_1^R - \phi^R$ for every $t < k$. Hence, in expectation, the SO's net transfers (i.e., commissions minus entry fees) to a distributor who purchases a license in period t are weakly smaller under $R^\star$ than they are under R given that σ is played. Thus, if σ is played in $\Gamma(R^\star)$, then the SO's expected profit will be (weakly) higher than the SO's expected profit when σ is played in $\Gamma(R)$, which is $\pi(R)$.

Since $qv_k b_1^{R^\star} = c$, every agent who enters $\Gamma(R^\star)$ in each period $t \leq k$ (respectively, $t \geq k$) finds it optimal to purchase (respectively, not to purchase) a license in every SPE of $\Gamma(R^\star)$. Since each distributor who interacts with an agent in $\Gamma(R^\star)$ is indifferent between selling a license to him and not doing so, σ is an SPE of[19] $\Gamma(R^\star)$.

Step 2. Let $\mathcal{Q}$ be the (finite) set of IC schemes such that $\frac{b_{\tau+1}^R}{b_\tau^R} = p$ for each $\tau \geq 1$, $b_1^R = \frac{c}{qv_t}$ for some $t \in \{1, ..., n-1\}$, $\phi^R = 0$, $\eta^R = 1$, and $a_\tau^R = 0$ for each $\tau \geq 1$. Step 1 showed that if the SO's expected profit is not maximized by selling directly to the agents, then $\mathcal{Q} \neq \emptyset$ and $max_{R \in \mathcal{Q}} \pi(R) \geq \pi(R')$ for every IC scheme R'.

Proof of Proposition 3. We shall need additional notation for this proof. 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 $\pi(R^n)$, where $\beta^n = (\beta_1^n, \beta_2^n)$. We use k^n to denote the last period in which the agents accept offers to purchase a license in σ^n. If such a period does not exist, then $k^n = 0$.

For each $t \geq 1$ and $j \in D_t$, if j sells a license to the t-th entrant, then we form a link $j \to i_t$. This induces a *distribution tree* T, rooted at the SO, where each node of the tree (besides the root) represents a distributor. Denote the set of rooted directed trees by $\mathbb{T}$ and let T^n be the random tree induced by σ^n. We use $Pr(T^n = T|\sigma^n)$ to

[19]The assumption that the distributors make offers when they are indifferent selects the SPE that induces the maximal expected profit for the SO but not necessarily σ.

denote the probability that σ^n results in the distribution tree T. For each $T \in \mathbb{T}$, the length of the directed path connecting i and j in T is denoted by $d_T(i,j)$.

Let us introduce some basic facts about R^n and (σ^n, β^n). Since the likelihood of meeting new entrants decreases as time progresses, every agent accepts every offer that he receives prior to period k^n. Since rejected offers only worsen the agents' expectations, every agent who receives an offer at $t = k^n$ accepts it and the SO does not make offers after period k^n. Since $p = 0$, every distributor makes an offer to every agent with whom he interacts and, thus, $\beta_2^n = 1$. The number of offers that each distributor makes after period k^n is $\sum_{t=k^n+1}^{n} \frac{1}{t}$. Since no offers are rejected prior to period k^n, the proportion of accepted offers is $\beta_1^n = \frac{1}{1+\sum_{j=k^n+1}^{n} \frac{1}{j}}$. Note that $\eta^{R^n} = 1$ as, otherwise, the SO's expected profit can be increased by raising η^{R^n}.

Lemmata 2 and 3 provide lower bounds for $\pi(R^n)$ and k^n, respectively.

Lemma 2 *There exist two numbers $\hat{\gamma} \in (0,1)$ and $n(\hat{\gamma}) \in \mathbb{N}$ such that $\pi(R^n) \geq \hat{\gamma} n$ for every $n > n(\hat{\gamma})$.*

Proof. Consider a small $\gamma > 0$ such that $\frac{c}{q(log(1/\gamma)-1)} < 1$ and set $k = \lceil \gamma n \rceil$. Let R be a scheme such that $\eta^R = 1$, $a_\tau^R = 0$ for every $\tau \geq 1$, $b_1^R = \frac{c}{q\sum_{i=k+1}^{n} \frac{1}{i}}$, $\phi^R = 0$, and $b_\tau^R = 0$ for every $\tau > 1$. By definition, $\pi(R^n) \geq \pi(R)$.

Consider a profile of strategies σ in which every agent $i \in I$ accepts (respectively, rejects) every offer that he receives at every time $t \leq k$ (respectively, $t > k$) and, for every $t \geq 1$, every $j \in D_t$ makes an offer if he meets an agent in period t.

As $p = 0$, the distributors' behavior in σ is optimal. Since $b_1^R = \frac{c}{q\sum_{i=k+1}^{n} \frac{1}{i}}$ and there are no other commissions, every agent finds it optimal to purchase a license if and only if $t \leq k$. The SO's expected profit is greater than $qk[1 + \sum_{i=k+1}^{n} \frac{1}{i}]\left(1 - b_1^R\right)$ under σ. If n is sufficiently large, then $b_1^R < \frac{c}{q(log(1/\gamma)-1)}$ and $1 + \sum_{i=k+1}^{n} \frac{1}{i} > log(1/\gamma)$. Hence,

$$\pi(R^n) \geq \pi(R) \geq n\hat{\gamma} := n\gamma q\left(1 - \frac{c}{q(log(1/\gamma)-1)}\right) log(1/\gamma) \tag{4}$$

■

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

Proof. Fix $k^n \geq 1$. The SO obtains at most $1 + B$ from each of the first k^n entrants and at most 1 from every agent who meets a distributor after period k^n. Hence,

$$\pi(R^n) \leq k^n(1+B) + (k^n+1)\sum_{i=k^n+1}^{n} \frac{1}{i} < k^n(2+B) + log(n) + k^n log(n/k^n) \tag{5}$$

Let $\bar{\gamma} \in (0,1)$ such that $\bar{\gamma}(2+B) < \hat{\gamma}/3$ and $\bar{\gamma} log(1/\bar{\gamma}) < min\{\hat{\gamma}/3, 1/e\}$. There exists $n(\bar{\gamma})$ such that if $n > n(\bar{\gamma})$ and $k^n \leq \bar{\gamma} n$, then the lower bound on the SO's expected profit that is obtained in (4) is greater than the upper bound in (5). ■

As $p = 0$, under σ^n, the SO makes offers to every agent that he meets in a set of adjacent periods $\{s^n, s^n + 1, ..., k^n\}$. Lemma 4 provides an upper bound for s^n, which denotes the first period in which the SO makes an offer under σ^n.

Lemma 4 *There exists a number $n^\star$ such that $s^n \leq \frac{2(1+B)}{\hat{\gamma}}$ for every $n > n^\star$.*

Proof. The SO's expected revenue from the retail sales made prior to period s^n is $q\sum_{i=1}^{s^n-1} \frac{1}{i} < q\left(log\left(n\right) + 1\right)$. Since $p = 0$, only the agents who will meet the SO after period $s^n - 1$ and their successors can buy a license (or purchase the good) from period s^n onward. The number of these potential buyers is the number of nodes in a subtree of a uniform random recursive tree of size $n + 1$ (excluding the root), rooted at the s^n-th node, which is, in expectation, $\frac{n+1-s^n}{s^n}$. Thus, the SO's expected revenue cannot exceed $qlog\left(n\right) + q + \frac{n+1-s^n}{s^n}\left(1 + B\right) \leq qlog\left(n\right) + q + \frac{n}{s^n}\left(1 + B\right)$. By Lemma 2, for large values of n, it must be that $\frac{n}{s^n}\left(1 + B\right) \geq \frac{\hat{\gamma}n}{2}$ and, therefore, $s^n \leq \frac{2(1+B)}{\hat{\gamma}}$. ■

Lemma 5 examines the induced number of distributors at the optimum, $|T^n| - 1$.

Lemma 5 *Fix arbitrary $\epsilon > 0$ and $\rho > 0$. There exists a number $n\left(\epsilon, \rho\right)$ such that $Pr\left(|T^n| > \rho\right) > 1 - \epsilon$ for every $n > n\left(\epsilon, \rho\right)$.*

Proof. Note that $|T^n|$ is the number of nodes in a subtree of a uniform random recursive tree of size $k^n + 1$, rooted at the s^n-th node. By Theorem 1 in Athreya and Ney (1972, p. 220), as k^n goes to infinity, the limiting distribution of $\frac{|T^n|}{k^n+1}$ is $Beta\left(1, s^n - 1\right)$. By Lemma 3, $k^n > \bar{\gamma}n$ for sufficiently large values of n. By Lemma 4, s^n is bounded from above. Thus, for a sufficiently large n, $Pr\left(|T^n| > \rho\right)$ is arbitrarily close to 1. ■

We now examine the SO's "dual" problem: finding an IC scheme R^n that minimizes his expected cost given ϕ^{R^n}, η^{R^n}, and that is subject to the requirement that σ^n be part of an ABEE of $\Gamma\left(R^n\right)$. Denote by κ_z the increase in the SO's expected costs (given that σ is being played) when the commission z is increased. The dual problem is:

$$min_{a_1,...,a_n,b_1,...,b_n} \kappa_{a_1}a_1 + ... + \kappa_{a_n}a_n + \kappa_{b_1}b_1 + ... + \kappa_{b_n}b_n \tag{6}$$

$$s.t \ (i) \ \sigma^n \text{ is an ABEE of the induced game} \tag{7}$$

$$(ii) \ a_\tau \leq \phi^{R^n} \text{ and } b_\tau \leq 1 \text{ for every } \tau \in \{1, ..., n\} \tag{8}$$

Since $p = 0$, the distributors' behavior in σ^n is optimal regardless of the scheme that is used. Thus, if the k^n-th entrant is indifferent between buying a license and not, then (7) holds. If (7) holds and the k^n-th entrant strictly prefers buying a license, we can reduce the SO's costs by scaling down the commissions. Thus, we can write (7) as:

$$w_{a_1}a_1 + ... + w_{a_{n-k^n}}a_{n-k^n} + w_{b_1}b_1 + ... + w_{b_{n-k^n}} = \phi^{R^n} + c \tag{9}$$

where w_z is the increase in the k^n-th entrant's willingness to pay for a license when the commission z is increased and the ABEE (σ^n, β^n) is played.

Lemma 6 *For every two commissions $a_\tau^{R^n}$ and $z \in \{a_1, ..., a_{n-k^n}, b_1, ..., b_{n-k^n}\}$ (respectively, $b_\tau^{R^n}$ and $z \in \{a_1, ..., a_{n-k^n}, b_1, ..., b_{n-k^n}\}$) such that $\tau \leq n - k^n$, if $z > 0$ and $\frac{\kappa_{a_\tau}}{w_{a_\tau}} < \frac{\kappa_z}{w_z}$ (respectively, $\frac{\kappa_{b_\tau}}{w_{b_\tau}} < \frac{\kappa_z}{w_z}$), it is the case that $a_\tau^{R^n} = \phi^{R^n}$ (respectively, $b_\tau^{R^n} = 1$).*

Proof. The proof follows directly from the linearity of (6) and (9). ∎

We now examine the costs of the recruitment commissions. Let $\mathbb{1}\left(d_T\left(SO, j\right) > \tau\right) \in \{0, 1\}$ be an indicator that equals 1 if and only if $d_T\left(SO, j\right) > \tau$.

Lemma 7 *Fix $\epsilon > 0$ and $\tau \geq 0$. There exists a number $n_{\epsilon,\tau}$ such that if $n > n_{\epsilon,\tau}$, then*

$$\kappa_{a_\tau} = \sum_{T \in \mathbb{T}} \sum_{j \in T} Pr\left(T^n = T | \sigma^n\right) \mathbb{1}\left(d_T\left(SO, j\right) > \tau\right) > (1 - \epsilon) \sum_{T \in \mathbb{T}} \sum_{j \in T} Pr\left(T^n = T | \sigma^n\right)$$

Proof. The tree T^n is a uniform random recursive tree rooted at the SO. Denote the i-th agent to buy a license by j. The distance $d_T\left(SO, j\right)$ corresponds to the distance between the root and the $(i+1)$-th node of a uniform random recursive tree. It is referred to as the *insertion depth* of the $(i+1)$-th node. A central limit theorem (Theorem 1 in Mahmoud, 1991) shows that the normalized insertion depth $M_i^\star = \frac{M_i - log(i)}{\sqrt{log(i)}}$ has the limiting distribution $\mathcal{N}(0, 1)$, the standard normal distribution. Hence, for every $\epsilon' > 0$ and $\tau \geq 0$, there exists $\rho_{\epsilon',\tau}$ such that for every $\rho > \rho_{\epsilon',\tau}$:

$$\sum_{T:|T| = \rho} \sum_{j \in T} \frac{\mathbb{1}\left(d_T\left(SO, j\right) > \tau\right)}{(\rho - 1)!} > (1 - \epsilon') \rho \tag{10}$$

Since the agents are equally likely to meet new entrants, for every $T \in \mathbb{T}$, it holds that $Pr\left(T^n = T | \sigma^n\right) = \frac{1}{(|T|-1)!} Pr\left(|T^n| = |T| | \sigma^n\right)$. We can write the premise as:

$$\kappa_{a_\tau} = \sum_{\rho=1}^{k_n - s_n + 2} \sum_{T:|T| = \rho} \sum_{j \in T} \frac{Pr\left(|T^n| = \rho | \sigma^n\right)}{(\rho - 1)!} \mathbb{1}\left(d_T\left(SO, j\right)\right) > (1 - \epsilon) \sum_{\rho=1}^{k_n - s_n + 2} \rho Pr\left(|T^n| = \rho | \sigma^n\right)$$

By (10), the above inequality holds for large values of ρ. By Lemma 5, for any $\rho \in \mathbb{N}$, $\epsilon' > 0$, and sufficiently large value of n, it holds that $Pr\left(|T^n| < \rho | \sigma^n\right) < \epsilon'$. It follows that the above inequality holds for sufficiently large values of n. ∎

When a distributor j such that $d_T\left(SO, j\right) = \tau$ makes a sale, the SO pays a total of $b_1^{R^n} + ... + b_\tau^{R^n}$, and if j purchases the good, then the SO pays a total of $b_1^{R^n} + ... + b_{\tau-1}^{R^n}$. Hence, $\kappa_{b_\tau} = q\left(\kappa_{a_\tau} + \kappa_{a_{\tau-1}} \sum_{t=k^n+1}^{n} \frac{1}{t}\right)$ where $q \sum_{t=k^n+1}^{n} \frac{1}{t}$ is the expected number of sales per distributor after period k^n.

Denote by l_τ^n the expected number of agents at the τ-th level of the subtree of G, rooted at the k^n-th entrant (e.g., $l_1^n = \sum_{j=k^n+1}^{n} \frac{1}{j}$). Note that $\beta_1^n = \frac{1}{1 + l_1^n}$. Since $p = 0$

and $\beta_2^n = 1$, the k^n-th entrant analogy-based expects to have $l_\tau^n (\beta_1^n)^\tau$ distributors at the τ-th level of his downline. Thus, $w_{a_\tau} := l_\tau^n (\beta_1^n)^\tau$ and $w_{b_\tau} = q l_\tau^n (\beta_1^n)^{\tau-1}$.

Lemma 8 is a technical result on random trees and its proof appears in the Technical Results section of Appendix B.

Lemma 8 *For every n and $\tau \in \{1, ..., n-1\}$ it holds that $l_1^n l_\tau^n \geq 2 l_{\tau+1}^n$.*

Lemma 9 studies the structure of R^n for large values of n using the ratio $\frac{\kappa}{w}$.

Lemma 9 *Fix an arbitrary integer $\tau^\star > 1$. There exists a number $n_{\tau^\star}$ such that $a_\tau^{R^n} = 0$ and $b_\tau^{R^n} = 0$ for every pair (n, τ) such that $n > n_{\tau^\star}$ and $\tau \in \{2, ..., \tau^\star\}$.*

Proof. By Lemma 8, $l_1^n l_\tau^n \geq 2 l_{\tau+1}^n$ for every n and $\tau \in \{1, ..., n-1\}$. Hence, $w_{a_\tau} \geq 2 w_{a_{\tau+1}}$ and $w_{b_\tau} \geq 2 w_{b_{\tau+1}}$. Note that $q w_{a_\tau} = w_{b_\tau} \beta_1^n$ for every n and $\tau \leq n - k_n$. By Lemma 7, for every $\tau^\star \in \mathbb{N}$, there exists a number $n_{\tau^\star}$ such that $\frac{\kappa_{b_\tau}}{\kappa_{b_{\tau+1}}}$ and $\frac{\kappa_{a_\tau}}{\kappa_{a_{\tau+1}}}$ are arbitrarily close to 1, and $\frac{\kappa_{b_\tau}}{\kappa_{a_\tau}}$ is arbitrarily close to $q(1 + l_1^n)$ for every pair $\tau \leq \tau^\star$ and $n > n_{\tau^\star}$. Hence, for every pair $\tau \leq \tau^\star$ and $n > n_{\tau^\star}$, it holds that

$$\frac{\kappa_{a_\tau}}{w_{a_\tau}} < \frac{\kappa_{a_{\tau+1}}}{w_{a_{\tau+1}}}, \frac{\kappa_{b_\tau}}{w_{b_\tau}} < \frac{\kappa_{b_{\tau+1}}}{w_{b_{\tau+1}}}, \frac{\kappa_{a_\tau}}{w_{a_\tau}} < \frac{\kappa_{b_{\tau+1}}}{w_{b_{\tau+1}}}, \text{ and } \frac{\kappa_{b_\tau}}{w_{b_\tau}} < \frac{\kappa_{a_{\tau+1}}}{w_{a_{\tau+1}}} \tag{11}$$

By Lemma 6, the ratios in (11) imply that for every $n > n_{\tau^\star}$, if $a_\tau^{R^n} > 0$ or $b_\tau^{R^n} > 0$ for $\tau \in \{2, ..., \tau^\star\}$, then $a_1^{R^n} = \phi^{R^n}$ and $b_1^{R^n} = 1$. If $a_1^{R_n} = \phi^{R_n}$ and $b_1^{R_n} = 1$, then the SO's profit cannot exceed his own sales and recruitments. Thus, $\pi(R^n)$ cannot exceed $\sum_{j=1}^n \frac{(1+B)}{j}$. For large values of n, $\sum_{j=1}^n \frac{(1+B)}{j}$ is less than the lower bound for the SO's expected profit, given in (4). Hence, for every $\tau^\star \in \mathbb{N}$, there exists $n_{\tau^\star}$ such that $a_\tau^{R^n} = 0$ and $b_\tau^{R^n} = 0$ for each pair (n, τ) such that $n > n_{\tau^\star}$ and $\tau \in \{2, ..., \tau^\star\}$. ■

Lemma 10 provides an upper bound for $\sum_{\tau > \tau^\star} (w_{a_\tau} + w_{b_\tau})$ for sufficiently large values of $\tau^\star$ and n.

Lemma 10 *For every $\epsilon > 0$, there exist $\tau^\star \in \mathbb{N}$ and $n(\epsilon, \tau^\star)$ such that $\sum_{\tau > \tau^\star} l_\tau^n (\beta_1^n)^{\tau-1}$ $< \epsilon$ for every $n > n(\epsilon, \tau^\star)$.*

Proof. By Lemma 3, for sufficiently large n, it must be that $k^n > \bar{\gamma} n$ and, therefore, $l_1^n < log(1/\bar{\gamma}) + 1$. By Lemma 8, $l_1^n l_{\tau-1}^n \geq 2 l_\tau^n$ for every $\tau \leq n-1$. It follows that

$$\sum_{\tau > \tau^\star} \frac{l_\tau^n}{(1 + l_1^n)^{\tau-1}} < 0.5 (log(1/\bar{\gamma}) + 1) \sum_{\tau > \tau^\star} \frac{l_{\tau-1}^n}{(1 + l_1^n)^{\tau-1}}$$

Since $l_1^n l_{\tau-1}^n \geq 2 l_\tau^n$, it must be that $(log(1/\bar{\gamma}) + 1) \sum_{\tau > \tau^\star} \frac{l_{\tau-1}^n}{(1+l_1^n)^{\tau-1}} < \epsilon$ for large $\tau^\star$. ■

Consider large values of $\tau^\star$ and $n > n_{\tau^\star}$, and assume that $\phi^{R_n} > 0$. We shall show that the SO can increase his expected profit by means of a scheme that does not charge entry fees. Lemma 9 showed that $a_\tau^{R^n} = 0$ and $b_\tau^{R^n} = 0$, for every $\tau \in \{2, ..., \tau^\star\}$. Lemma

10 showed that, if $\tau^\star$ is sufficiently large, then, ex ante, a reward of at most $\epsilon\left(\phi^{R^n}+1\right)$ can be provided to the k^n-th entrant by means of the commissions $\left\{b_\tau^{R^n}:\tau>\tau^\star\right\}$ and $\left\{a_\tau^{R^n}:\tau>\tau^\star\right\}$, where $\epsilon>0$ can be chosen to be arbitrarily close to 0.

Let us eliminate the recruitment commissions (i.e., $a_1^{R^n}$ and $\left\{a_\tau^{R^n}:\tau>\tau^\star\right\}$) and the entry fees. The SO's expected revenue is reduced by ϕ^{R^n} multiplied by the expected number of distributors. By the inequality in the premise of Lemma 7, for a sufficiently large n, the SO's expected revenue is reduced by no more than $\phi^{R^n}\frac{\kappa_{a_1}}{1-\epsilon}$. In addition to this change, let us reduce $b_1^{R^n}$ by $\phi^{R^n}\frac{\epsilon}{1-\epsilon}\frac{w_{a_1}}{w_{b_1}}+\left(\phi^{R^n}-a_1^{R_n}\right)\frac{w_{a_1}}{w_{b_1}}$. The total expected reduction in the SO's cost is greater than

$$a_1^{R^n}\kappa_{a_1}+\phi^{R^n}\frac{\epsilon}{1-\epsilon}\frac{w_{a_1}}{w_{b_1}}\kappa_{b_1}+\left(\phi^{R^n}-a_1^{R^n}\right)\frac{w_{a_1}}{w_{b_1}}\kappa_{b_1} \tag{12}$$

Since $\frac{\kappa_{a_1}}{w_{a_1}}=\frac{\kappa_{a_1}}{l_1^n\beta_1^n}=\frac{\kappa_{a_1}+\kappa_{a_1}l_1^n}{l_1^n}<\frac{q\kappa_{a_1}+q\kappa_{a_0}l_1^n}{ql_1^n}=\frac{\kappa_{b_1}}{w_{b_1}}$, (12) must be greater than $\phi^{R^n}\frac{\kappa_{a_1}}{1-\epsilon}$.

Eliminating the recruitment commissions reduces the k^n-th entrant's willingness to pay for a license by at most $a_1^{R^n}w_{a_1}+\phi^{R^n}\epsilon$. The reduction in $b_1^{R^n}$ reduces his willingness to pay for a license by at most $\left(\phi^{R^n}-a_1^{R^n}\right)w_{a_1}+\phi^{R_n}\frac{\epsilon}{1-\epsilon}w_{a_1}$. By Lemma 3, $l_1^n<log\left(1/\bar{\gamma}\right)+1$ such that w_{a_1} is bounded below 1. Thus, for sufficiently small ϵ, the reduction in the k^n-th entrant's willingness to pay for a license is less than ϕ^{R^n}. This contradicts $\phi^{R_n}>0$ being part of a profit-maximizing scheme as the change renders both the SO and the k^n-th entrant better off.

It is left to verify that $b_1^{R^n}\geq\phi^{R^n}\frac{\epsilon}{1-\epsilon}\frac{w_{a_1}}{w_{b_1}}+\left(\phi^{R^n}-a_1^{R^n}\right)\frac{w_{a_1}}{w_{b_1}}$ such that the above exercise is viable. If this inequality does not hold, then the k^n-th entrant's willingness to pay for a license cannot exceed

$$\phi^{R^n}\frac{\epsilon}{1-\epsilon}\frac{w_{a_1}}{w_{b_1}}w_{b_1}+\left(\phi^{R^n}-a_1^{R^n}\right)\frac{w_{a_1}}{w_{b_1}}w_{b_1}+\epsilon\phi^{R^n}+\epsilon+a_1^{R^n}w_{a_1} \tag{13}$$

Since $\phi^{R^n}\leq B$, if $\epsilon>0$ is sufficiently small, then (13) is less than $\phi^{R^n}+c$, which is in contradiction to the optimality of the k^n-th entrant's decision to purchase a license.

In conclusion, for large values of n, it holds that $\phi^{R^n}=0$, for otherwise, the SO could increase his expected profit by means of an IC scheme that does not charge entry fees. Since R^n is IC and $\phi^{R^n}=0$, it follows that $a_\tau^{R^n}=0$ for every $\tau\geq 1$.

Proof of Proposition 4. Consider an agent j who contemplates purchasing a license in period $t\in\mathbb{N}$. He expects to have $\sum_{i=1}^{\infty}\frac{\delta_j^i}{t+1}\leq\sum_{i=1}^{\infty}\frac{\bar{\delta}^i}{t+1}$ successors. If R is IC, j expects a payoff smaller than $\left(\phi^R+q\right)\sum_{i=1}^{\infty}\frac{\bar{\delta}^i}{t+1}-\phi^R-c$ if he purchases a license. If $\phi^R>0$ or $c>0$, then there is a period $t^\star$ from which point onward purchasing a license is strictly suboptimal regardless of j's beliefs about his successors' behavior.

A standard backward induction argument shows that, in every SPE of a game that is induced by any IC scheme, the agents never purchase a license if $\phi^R > 0$.

Appendix B

A Semistationary Model

Our main objective is to show that the paper's main insights do not depend on the finiteness of the game. We shall focus on Theorems 1–4 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, as in the main text, in each period $t \in \mathbb{N}$, nature draws one agent $i \in I$ to enter the game, uniformly at random. In order to facilitate 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 at t is $\bar{\sigma}_t := \int_{j \in I} \sigma_j(t)\, dj$, where $\sigma_j(t) = 0$ (respectively, $\sigma_j(t) = 1$) if j rejects (respectively, accepts) offers he receives in period t. Let $r_\sigma(t)$ be the 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 w.r.t. β. The rest of the modeling assumptions remain as in the main text.

Propositions 5, 6, and 7 are analogous to Theorem 1, Proposition 2 (from which Theorems 2.1 and 3 follow), and Theorem 4, respectively. The proofs of these results are similar to the proofs of the results in the main text except for one main difference: 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.

Proposition 5 *Let $q = 0$. There exists no IC one-level scheme R such that $\pi(R) > 0$.*

Proof. Let R be an IC one-level scheme such that $\phi^R > 0$. Assume by way of contradiction that there exists an ABEE of $\Gamma(R)$ in which the agents purchase licenses.

Since the agents are equally likely to meet new entrants, in expectation, an agent who enters the game in period k will have $S_k = \sum_{j=1}^{\infty} \frac{\delta^j}{k+1}$ successors in G. He cannot analogy-based expect to sell more than $S_k\beta_1$ licenses if he purchases one and, as R is IC, he cannot analogy-based expect a reward greater than $S_k\beta_1 a_1^R - \phi^R \leq S_k\beta_1\phi^R - \phi^R$ if he purchases a license. A distributor who interacts with the k-th entrant cannot analogy-based expect to sell more than $pS_k\beta_1$ licenses to the k-th entrant's successors if he refrains from selling a license to the k-th entrant. As $lim_{k\to\infty}S_k = 0$, there exists a period t such that, in every ABEE of $\Gamma(R)$, *after* period t every agent rejects every offer and every distributor makes an offer to every agent with whom he interacts.

We shall prove Proposition 5 by induction on the size of t. We shall show that if no agent purchases a license after period t and every distributor makes an offer to every agent with whom he interacts after period t, then, in period t, no agent purchases a license and every distributor makes an offer to every agent with whom he interacts.

The LHS (respectively, RHS) of (14) is the expected (respectively, analogy-based expected) number of offers that a distributor who buys a license in period t makes.

$$v_t = \sum_{j=t+1}^{\infty} \frac{\delta^{j-t}}{j} + p \sum_{j=t+1}^{\infty} \sum_{j'=j+1}^{\infty} \frac{\delta^{j'-t}}{jj'} + p^2 \sum_{j=t+1}^{\infty} \sum_{j'=j+1}^{\infty} \sum_{j''=j+1}^{\infty} \frac{\delta^{j''-t}}{jj'j''} + ... \geq \tag{14}$$

$$\hat{v}_t = \sum_{j=t+1}^{\infty} \frac{\delta^{j-t}}{j} + p(1-\beta_1) \sum_{j=t+1}^{\infty} \sum_{j'=j+1}^{\infty} \frac{\delta^{j'-t}}{jj'} + p^2 (1-\beta_1)^2 \sum_{j=t+1}^{\infty} \sum_{j'=j+1}^{\infty} \sum_{j''=j+1}^{\infty} \frac{\delta^{j''-t}}{jj'j''} + ...$$

Each offer that is accepted up to period t results in a distributor who, if the game reaches period t, in expectation, makes v_t offers to the agents that he meets after period t and their successors. Thus, $\beta_1 \leq \frac{1}{1+\delta^t v_t}$. It follows that $\hat{v}_t\beta_1 \leq \frac{v_t}{1+\delta^t v_t} \leq \frac{\frac{\delta}{(t+1)(1-\delta)}}{1+\delta^t \frac{\delta}{(t+1)(1-\delta)}} < 1$, where the second inequality results from plugging $p = 1$ into v_t.

Since R is IC, $\phi^R \geq a_1^R$. In an ABEE, the agents do not purchase a license in period t as $\hat{v}_t\beta_1\phi^R - \phi^R < 0$. Every distributor makes an offer to every agent with whom he interacts in period t as he analogy-based expects to sell $p\hat{v}_t\beta_1 < 1$ licenses to the t-th entrant's successors if he does not sell a license to him. By induction, we obtain a contradiction to the existence of an ABEE in which the agents purchase licenses when $\phi^R > 0$. As $q = 0$, there exists no IC scheme R such that $\pi(R) > 0$. ■

Proposition 6 *Fix $q = 0$. There exists a number $\delta^\star < 1$ such that for every $\delta > \delta^\star$ there exists a number $p^\star(\delta)$ such that if $p \leq p^\star(\delta)$, then there exists an IC two-level scheme R such that $\pi(R) > 0$.*

Proof. The proof is identical to the proof of Proposition 2 for the case of $\alpha = 1$ except for three changes: the analogy-based expectations change from $\beta_1 = \frac{1}{1+\sum_{t=2}^{n}\frac{1}{t}}$ to $\hat{\beta}_1 = \frac{1}{1+\sum_{i=1}^{\infty}\frac{\delta^i}{1+i}}$, the number of agents in the first level of the subtree of G rooted at the first entrant changes from $l_1 = \sum_{t=2}^{n}\frac{1}{t}$ to $\hat{l}_1 = \sum_{i=1}^{\infty}\frac{\delta^i}{1+i}$, and the number of agents at the second level of that subtree changes from $l_2 = \sum_{t=2}^{n-1}\sum_{t'=t+1}^{n}\frac{1}{tt'}$ to $\hat{l}_2 = \sum_{i=1}^{\infty}\sum_{i'=i+1}^{\infty}\delta^{i'}\frac{1}{(1+i)(1+i')}$.

As $\lim_{\delta\to 1}\hat{\beta}_1\hat{l}_1 = 1 = \lim_{n\to\infty}\beta_1 l_1$ and $\lim_{\delta\to 1}\hat{\beta}_1^2\hat{l}_1 = \frac{1}{2} = \lim_{n\to\infty}\beta_1^2 l_2$, these changes do not affect the proof. For $\delta < 1$ sufficiently close to 1, the first entrant analogy-based expects a reward arbitrarily close to $\frac{3}{2}\phi^R x - \phi^R - c$, where $x < 1$ and ϕ^R can be set such that this expression is equal to 0. ∎

Proposition 7 *Let $c > 0$ and $q > 0$. There exists an IC profit-maximizing scheme $R^\star$ such that $\phi^{R^\star} = 0$ and $a_\tau^{R^\star} = 0$ for every $\tau \geq 1$.*

Proof. Consider an arbitrary IC scheme R and an agent j who contemplates purchasing a license in period $t \in \mathbb{N}$. He expects to have $\sum_{i=1}^{\infty}\frac{\delta^i}{t+1}$ successors. Agent j expects a payoff smaller than $\left(\phi^R + q\right)\sum_{i=1}^{\infty}\frac{\delta^i}{t+1} - \phi^R - c$ in case he purchases a license. As $c > 0$, there is a period $t^\star$ from which point onward purchasing a license is strictly suboptimal regardless of j's beliefs about his successors' behavior. Hence, there exists a period $t^\star$ such that for every $t > t^\star$, every IC scheme R, and every agent $j \in I$, purchasing a license in period t of $\Gamma(R)$ is strictly suboptimal for j regardless of his beliefs about the other agents' strategies. The rest of the proof is identical to the proof of Theorem 4 except for two changes. First, we need to replace n with $t^\star$ in the last paragraph. Second, the expression for v_t changes to the LHS of (14) instead of (1). ∎

Technical Results

Theorem 2: Proof that the derivative of (3) w.r.t. p is non-positive.

The derivative of the z-th component of (3), given by $\frac{l_z(1+pv_k(p))^{z-1}}{(1+v_k(p))^z}$, w.r.t. p is:

$$\frac{l_z}{(1+v_k(p))^{2z}}\left[(z-1)(1+v_k(p))^z(1+pv_k(p))^{z-2}(v_k(p)+v_k'(p)p)\right] - \frac{l_z}{(1+v_k(p))^{2z}}\left[z(1+v_k(p))^{z-1}(1+pv_k(p))^{z-1}v_k'(p)\right] \quad (15)$$

The derivative of (3) w.r.t. p can be written as $\sum_{z=1}^{n-k}\frac{z(1+pv_k(p))^{z-1}}{(1+v_k(p))^{z+1}}[l_{z+1}(v_k(p)+v_k'(p)p)-$

$l_z v_k'(p)]$ since $l_{n-k+1} = 0$. It equals the sum of the following square matrix's elements:

$$\begin{pmatrix} (l_2 l_1 - l_1 l_2)\gamma_{11} & (l_2 l_2 - l_1 l_3)\gamma_{12} & (l_2 l_3 - l_1 l_4)\gamma_{13} & \dots & (l_2 l_{n-k} - 0)\gamma_{1(n-k)} \\ (l_3 l_1 - l_2 l_2)\gamma_{21} & 0 \times \gamma_{22} & (l_3 l_3 - l_2 l_4)\gamma_{23} & \dots & \vdots \\ (l_4 l_1 - l_3 l_2)\gamma_{31} & (l_4 l_2 - l_3 l_3)\gamma_{32} & 0 \times \gamma_{33} & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (0 - l_{n-k} l_2)\gamma_{(n-k)1} & \dots & \dots & \dots & 0 \end{pmatrix}$$

where $\gamma_{xy} = \frac{xyp^{y-1}(1+pv_k(p))^{x-1}}{(1+v_k(p))^{x+1}}$. Observe that $\gamma_{xy} \geq \gamma_{yx}$ if and only if $x \geq y$. Hence, the derivative of (3) w.r.t. p is (weakly) negative if $\frac{l_2}{l_1} \geq \frac{l_3}{l_2} \geq ... \geq \frac{l_{n-k}}{l_{n-k-1}}$.

To prove the above inequality, consider the subtree of G rooted at the k-th entrant and denote by $o_{\lambda,t}$ the expected number of agents at the λ-th level of that subtree at the end of period $t > k$. An increase in t strictly raises $\frac{o_{2,t}}{o_{1,t}}$ since $o_{1,t} - o_{1,t-1} = \frac{1}{t}$ and $o_{2,t} - o_{2,t-1} = \frac{1}{t}\left(\frac{1}{k+1} + ... + \frac{1}{t-1}\right)$. The LHS of (16) is weakly increasing in t as $\frac{o_{2,t}}{o_{1,t}}$ is increasing in t.

$$\frac{o_{\lambda+1,t}}{o_{\lambda,t}} = \frac{o_{\lambda+1,t-1} + \frac{1}{t} o_{\lambda,t-1}}{o_{\lambda,t-1} + \frac{1}{t} o_{\lambda-1,t-1}} \tag{16}$$

Extend the branching process to $n+1$ periods and consider $\lambda \leq n-k$ such that (16) is well defined for $t = n+1$. Since the LHS of (16) is weakly increasing in t, it must be that $\frac{o_{\lambda+1,n+1}}{o_{\lambda,n+1}} \geq \frac{o_{\lambda+1,n}}{o_{\lambda,n}}$. The equality in (16) implies that $\frac{o_{\lambda,n}}{o_{\lambda-1,n}} \geq \frac{o_{\lambda+1,n+1}}{o_{\lambda,n+1}} \geq \frac{o_{\lambda+1,n}}{o_{\lambda,n}}$. Since $\frac{o_{\lambda,n}}{o_{\lambda-1,n}} = \frac{l_\lambda}{l_{\lambda-1}}$, it follows that $\frac{l_2}{l_1} \geq ... \geq \frac{l_{n-k}}{l_{n-k-1}} \geq \frac{l_{n-k+1}}{l_{n-k}} = 0$.

Proof of Lemma 8. Define $l^n_{\tau,t}$ to be the expected number of agents at the τ-th level of the subtree of G rooted at the t-th entrant. Note that $l^n_{\tau,t} = 0$ for every $\tau > n - t$. Observe that for any $t \leq n$ it holds that $\left(l^n_{1,t}\right)^2 = \left(\sum_{j=t+1}^n \frac{1}{j}\right)^2 = 2l_{2,t} + \sum_{j=t+1}^n \frac{1}{j^2}$. We prove the lemma by induction on the size of τ. Assume that $l^n_{1,t} l^n_{\tau-1,t} \geq 2l^n_{\tau,t}$ for every $t \leq n$. Let us show that $l^n_{1,t} l^n_{\tau,t} \geq 2l^n_{\tau+1,t}$. We can write this inequality as:

$$l^n_{1,t}\left(\frac{l^n_{\tau-1,t+1}}{t+1} + \frac{l^n_{\tau-1,t+2}}{t+2} + ...\right) \geq 2\left(\frac{l^n_{\tau,t+1}}{t+1} + \frac{l^n_{\tau,t+2}}{t+2} + ...\right) \tag{17}$$

By the induction hypothesis, (17) holds. Thus, $l^n_1 l^n_\tau = l^n_{1,k^n} l^n_{\tau,k^n} \geq 2l^n_{\tau+1,k^n} = 2l^n_{\tau+1}$.