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

Yair Antler

University of Essex

August 10, 2017

Abstract

We find conditions under which a scheme organizer is able to design a pyramid scam to exploit partially sophisticated agents. Further, we characterize the pyramid scams in which such agents would be willing to participate. Motivated by the growing discussion on the legitimacy of multilevel marketing schemes and their resemblance to pyramid scams, we highlight the key differences between the two types of schemes and characterize the optimal multilevel marketing scheme when agents are fully rational.

---

1 Introduction

In various instances when individuals fail to act in their best interest, they become vulnerable to manipulation and exploitation by others. Although policy makers and consumer protection agencies may try to protect individuals from engaging in contracts and transactions whose main purpose is to take advantage of such failures, it is not always easy to draw the boundary between legitimate and exploitative transactions and contracts. Are there any observable characteristics that would enable us to distinguish between the two types of contracts? Can we identify conditions under which individuals are more vulnerable to exploitation? How can regulators protect individuals while still allowing for legitimate business endeavors to thrive?

We address these and other questions in the context of the growing discussion[1] on the legitimacy of multilevel marketing companies (MLMs), and the problem of distinguishing between legitimate MLMs and pyramid scams.[2] A legitimate MLM is a company that uses independent representatives to sell genuine products or services to friends and acquaintances. These representatives earn commissions not only from the retail sales they make, but also from recruiting new representatives into their down-line as well as from the sales these new representatives make. According to the Direct Sellers Association, an MLM lobbying group, American MLMs made over 36 billion dollars in retail sales in 2015, and over 20 million Americans were involved in multilevel marketing in the past decade.[3]

In a pyramid scam, there is no genuine product and "representatives" are typically compensated for recruiting new members into their down-line. Pyramid scams are illegal in virtually every country and this provides their organizers with an incentive to disguise themselves as legitimate MLMs by promoting useless products (Federal Trade Commission, 2016). They are capable of doing so since, typically, legitimate MLMs sell products whose quality is hard to assess and verify, such as vitamins and nutritional supplements (Direct Selling Association, 2016; Schuette, n.d.). The difficulty in assessing the quality of such products is inherent in multilevel marketing since it is what makes a recommendation from a trusted acquaintance valuable and makes word-

---

[1]See, e.g., Bort (2016), Hoyle (2016), McCrum (2014), McKown (2017), McCarthy (2016), Nocera (2015), "Multi-level marketing in America" (2015), Parloff (2016), and Wieczner (2017).

[2]The Federal Trade Commission's website suggests that "there are multi-level marketing plans —and then there are pyramid schemes. Before signing on the dotted line, study the company's track record, ask lots of questions, and seek out independent opinions about the business" (Federal Trade Commission, 2016a).

[3]See Direct Selling Association (2016). Since MLMs are privately owned and sell different products, most of the available data is from lobby groups.

of-mouth marketing a viable tool to start with. The downside is that the difficulty in verifying the value of the merchandise is what makes it possible for pyramid sellers to disguise themselves as legitimate MLMs and makes it hard for consumer protection authorities and potential representatives to distinguish between these businesses (see Federal Trade Commission, 2016; Schuette, n.d.).

Our main objective is to identify both the differences between legitimate MLMs and pyramid scams, and the conditions that allow such scams to work. To this end, we develop a model of word-of-mouth marketing in which a scheme organizer (SO) distributes a good to a directed network of fully rational agents that is formed randomly and sequentially. These agents can purchase a distribution license from the SO or from other distributors and become distributors themselves. A distributor j is said to belong to another distributor i's down-line if he purchased a license from i or from another member of i's down-line. The distributors are compensated according to a reward scheme that maps their down-line and the retail sales made by members of their down-line to a monetary reward. The SO chooses the reward scheme while taking the agents' equilibrium behavior in its induced game into account.

The distributors are intermediaries in this setting. What makes intermediation valuable to the SO? In our model, agents must become aware of the existence of the good in order to purchase it. They learn of its existence from their network neighbors. With no financial incentives, these neighbors spread the information with some probability $p < 1$. The reward scheme provides distributors with financial incentives to tell their neighbors about the good such that the information about its existence (and the possibility of purchasing a distribution license) spreads to a larger number of agents.

In the real world, we find evidence of both legitimate multilevel marketing schemes and pyramid scams, the latter of which are prevalent despite being illegal.[4] While legitimate multilevel marketing schemes are used in order to distribute genuine goods, the main "product" being traded in a pyramid scam is the right to join the pyramid. Since fully rational economic agents *will never take part* in a pyramid scam, it is necessary to change our behavioral model in order to understand such scams and the extent to which they are different from legitimate multilevel marketing schemes.

In the contexts of speculation in financial markets and bubbles, which are closely related to pyramid scams, the literature offers two main approaches that depart from the rational expectations model: assuming that agents use *simplified representations of other agents' strategies/information* (Jehiel, 2005; Eyster and Rabin, 2005; Jehiel and

---

[4]For two examples of recent pyramid scams that lured a combined total of 406,000 "consumers" see Federal Trade Commission (2012, 2014).

Koessler, 2008; Eyster and Rabin, 2010; Bianchi and Jehiel, 2010; Eyster and Piccione, 2013; Steiner and Stewart, 2015) and relaxing the common prior assumption (Harisson and Kreps, 1978; Morris, 1996). We follow the former approach and use *analogy-based expectations equilibrium* (Jehiel, 2005) as a solution concept.[5]

We now illustrate the behavioral assumption underlying this solution concept in the context of our model. Consider a reward scheme and assume that it is commonly known that the willingness to pay for the good is 0 (i.e., the only "product" being traded is the right to sell distribution licenses). Agents receive offers to join the pyramid (i.e., to purchase a license) at different times. Roughly speaking, joining the pyramid may be more attractive to those agents who have an opportunity to join earlier since the number of pyramid members who compete for recruiting new entrants is smaller. Our agents understand this logic and have accurate beliefs regarding the process that generates the network. Using a simplified model of their opponents' strategies, they correctly grasp their opponents' average behavior but they fail to perceive the connection between this behavior and the time when their opponents make their decision to join the pyramid. Our agents can be viewed as if they were best responding to *statistically correct feedback* about other agents' behavior in *similar schemes* (e.g., the probability that an agent who is offered an opportunity to purchase a distribution license will purchase it) that is supplied by the SO.

We show that if the tendency of agents to mention the good to their network neighbors in the absence of financial incentives p is small, and the number of agents is sufficiently large, then the SO can construct a reward scheme that extracts gains from the agents even when there is no demand for the good. That is, *a pyramid scam is sustainable* in such a case. What type of reward schemes can systematically fool our partially sophisticated agents into taking part in such a pyramid scam? It turns out that there is a surprisingly simple property that characterizes these reward schemes: they must offer compensation that is based on *at least two layers* of an agent's down-line. We show that two layers of compensation are sufficient to construct a pyramid scam, whereas a scheme that pays commissions based on one layer of agents' down-lines cannot generate a strictly positive expected profit for the SO. This result suggests a rationale for why MLMs often offer recruitment-based compensation that is dependent on multiple levels of distributors' down-lines.

It is of interest to understand what individuals are especially vulnerable to pyramid scams and what their *degree of sophistication* is. To this end, we develop a cognitive hierarchy model in the spirit of Camerer, Ho, and Chong (2004). In our cognitive

---

[5]In Section 5, we analyze our model using the alternative approach offered by the literature.

hierarchy model, naivety (i.e., level 0's behavior) takes the form of *best responding to one's opponents' average behavior*. Observe that level 0's behavior corresponds to our notion of partial sophistication. This implies that, unlike in conventional level-k thinking models in which the naive type's behavior is exogenously given, here *the naive type's behavior is determined in equilibrium*. We show that, for every $k > 0$, an SO who faces a large population of agents whose level of sophistication is k *and above* can sustain a pyramid scam regardless of p. That is, agents whose degree of sophistication is greater than 0 may fall victim to a pyramid scam even in instances in which a population of level-0 agents would never participate in such a scam.

To understand the difference between pyramid scams and legitimate multilevel marketing, we solve for the profit-maximizing reward scheme when agents are fully rational and the SO produces a genuine product (i.e., the agents' willingness to pay for the good is strictly greater than 0). First, this scheme does not include entry fees. Second, it does not compensate distributors for recruiting new distributors into their down-line. Third, it pays commissions for multiple layers of retail sales (that is, for both direct and indirect retail sales). In other words, in the optimal scheme *only retail sales are rewarded*. Therefore, preventing MLMs from offering recruitment-based commissions and charging entry fees would have no effect on their profits.

Entry fees and recruitment-based compensation are crucial for real-world pyramid scams as there is no real product that can generate income for the SO.[6] At a glance, it seems that it is possible to prevent MLMs from offering recruitment-based pay or from charging entry fees. However, such a policy is hard to enforce: fees are sometimes disguised as costs of acquiring initial stock or training,[7] and the rewards are endogenous such that ex ante it is unclear what portion of them is related to sales and what portion is related to recruitment. In this paper, we attempt to obtain a better understanding of the mechanisms and incentives that underlie pyramid scams and multilevel marketing schemes in order to suggest a more transparent way of drawing the boundary between the two phenomena.

The paper proceeds as follows. We present the model in Section 2 and analyze legitimate multilevel marketing schemes in Section 3. Our main results are given in Section 4, which focuses on pyramid scams. In Section 5 we discuss alternative

---

[6]In some countries, consumer protection authorities define a pyramid scam as an MLM that pays "compensation for recruitment or requires purchases as a condition of participation"(Competition Bureau, 2015).

[7]"Pyramid companies make virtually all their profits from signing up new recruits and often attempt to disguise entry fees as the price charged for mandatory purchases of training, computer services, or product inventory" (Schuette, n.d.).

modeling approaches. Section 6 covers the related literature and Section 7 concludes.

2   The Model

There is a scheme organizer (SO) who produces a good free of cost and with no capacity constraints, and an infinite set of agents I. 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 decay parameter $\psi_i \in \{0, 1\}$. For every agent $i \in I$, we assume that ω_i and ψ_i are drawn independently and denote $p := Pr(\psi_i = 1)$ and $q := Pr(\omega_i = 1)$, respectively. The role and interpretation of the decay factor will be clarified soon.

Arrival process. Time $t = 0, 1, 2, ...$ is discrete. Conditional upon the game reaching period t, there is a probability $\delta < 1$ that the game continues and a probability $1 - \delta$ that it terminates at t. The SO enters the game at time 0. At each time $t > 0$, a new agent $j \in I$ enters the game and meets one player i who is chosen uniformly at random from the set of players who entered the game previously (including the SO). For example, an agent j who enters the game at time $t = 3$ meets the SO with probability $\frac{1}{3}$. Agent j meets the agent who entered at time 1 with probability $\frac{1}{3}$ and the agent who entered at time 2 with probability $\frac{1}{3}$. Let G denote the directed tree, rooted at the SO, that is induced by the meeting process. Observe that G evolves over time. An example of a snapshot of G at a specific point in time is given in Figure 1. The length of the directed path (if any exists) from i to j is denoted by $d_G(i, j)$ and the subtree of G rooted at $i \in I$ is denoted by G_i.
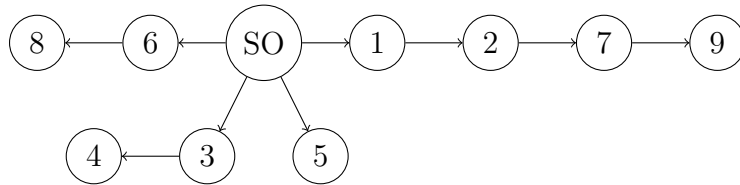

Figure 1: A snapshot of G at the end of period 9. The numbers represent the time at which each agent entered the game.

Distribution. For $t = 0, 1, ..., n$, let A_t denote the set of distributors at time t. Let $A_0 = \{SO\}$. We shall now describe the process of becoming a distributor. Each $i \in I$ who enters the game at a time $t \geq 1$ meets a player j and starts climbing the meeting tree G from j toward the SO. Agent i stops when he encounters an agent $k \notin A_{t-1}$ with $\psi_k = 0$ or a player $k' \in A_{t-1}$ (observe that k' may be the SO). In the former case, $A_t = A_{t-1}$ and the game moves to the next period. In the latter, i

interacts with k' as follows. Player k' can make agent i an offer to purchase a unit of the good at a fixed price η (predetermined by the SO), which is paid to the SO. He can also make agent i an offer to purchase a distribution license (k' is allowed to make both offers). Agent i may or may not agree to each of k''s offers. If agent i receives an offer to purchase a license and accepts it, then $A_t = A_{t-1} \cup \{i\}$. That is, i becomes a distributor. Otherwise, $A_t = A_{t-1}$. If i does not become a distributor in round t, he cannot become one later. That is, if $i \notin A_t$, then $i \notin A_\tau$ for all $\tau > t$.

We can now interpret the decay parameters p and ψ. Agent i's decay parameter ψ_i can be viewed as i's tendency to mention the product to people whom he meets when he does not have any financial incentive to do so. When an agent i (with $\psi_i = 1$) is not a distributor but has heard about the product from some other agent k, he will mention the product to agents whom he meets and refer them to agent k. The parameter p can be interpreted as a property of the good. When p is high, it means that the good raises interest and people tend to mention it when they talk to their friends and acquaintances.

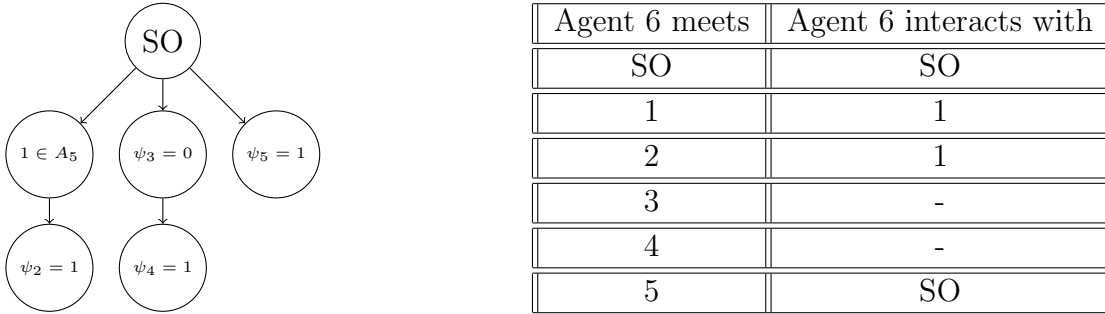| Agent 6 meets	Agent 6 interacts with
SO	SO
1	1
2	1
3	-
4	-
5	SO

Figure 2: A snapshot of the distribution process at the end of $t = 5$. On the left, we have the meeting tree G. The numbers represent the time at which each agent entered the game. Note that the set of distributors is $A_5 = \{SO, 1\}$. The possible scenarios for $t = 6$ are provided on the right. Recall that these scenarios are equally likely.

The distribution tree. When a distributor j offers a license to agent i and the latter accepts, a directed link $j \to i$ is formed and i becomes a *member of j's down-line*. We denote the number of units of the good sold by distributor j by w_j. Let T denote the directed weighted tree that reflects the sales in the SO's down-line at the end of the game: the skeleton of T is the SO's down-line and the weight of each node $i \in T$ is the number of sales made by distributor $i \in T$, that is, w_i. As before, we denote the length of the directed path from i to j in T by $d_T(i, j)$ if such a path exists. For each $i \in I$, we denote the subtree of T rooted at $i \in I$ by T_i.

Payoffs. An agent who purchases a distribution license and becomes a distributor

7

incurs a cost of $c > 0$ that reflects the effort of learning how to sell the good/licenses. Distributors receive commissions according to a reward scheme R chosen by the SO in advance. The reward scheme maps each distributor's down-line to a real reward. Formally, a reward scheme R sets a price η_R for the good[8] and pays each distributor i

$$\sum_{j \in T_i/\{i\}} a_R \left(d_T \left(i, j \right) \right) + \sum_{j \in T_i} b_R \left(d_T \left(i, j \right) + 1 \right) w_j - \phi_R \qquad (1)$$

where $\phi_R \geq 0$ is an entry fee paid to the SO, $a_R \left(1 \right), a_R \left(2 \right), \ldots \geq 0$ are recruitment commissions, and $b_R \left(1 \right), b_R \left(2 \right), \ldots \geq 0$ are retail sales commissions. Observe that this is with loss of generality. In Section 5, we discuss an extension of the model in which a reward scheme can be any function that maps directed weighted trees to real (possibly negative) rewards. An agent who purchases a unit of the good pays η_R to the SO and enjoys an additional payoff of ω_i.
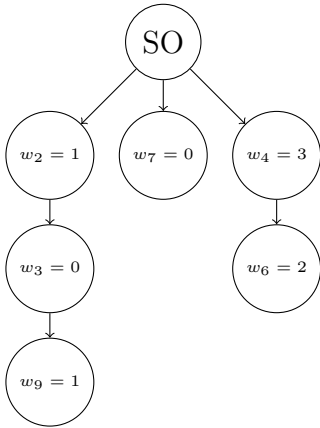
Distributor	Reward
2	$b_R \left(1 \right) + b_R \left(3 \right) + a_R \left(1 \right) + a_R \left(2 \right) - \phi_R$
3	$a_R \left(1 \right) + b_R \left(2 \right) - \phi_R$
4	$3 b_R \left(1 \right) + 2 b_R \left(2 \right) + a_R \left(1 \right) - \phi_R$
6	$2 b_R \left(1 \right) - \phi_R$
7	$-\phi_R$
9	$b_R \left(1 \right) - \phi_R$

Figure 3: An example of a distribution tree at the end of a game and its induced rewards. Observe that agents $1, 5$, and 8 are not distributors and therefore do not appear in the distribution tree T.

We restrict our attention to reward schemes in which $a_R \left(\tau \right) \leq \phi_R$ and $b_R \left(\tau \right) \leq \eta_R$ for every $\tau \in \mathbb{N}$, and refer to such schemes as *incentive-compatible* schemes. This restriction is motivated by the idea that the distributors have to report to the SO about the sales and recruitments they make. Schemes that are not incentive compatible provide each distributor with an incentive to misreport to the SO that he sold additional distribution licenses (or units of the good). By misreporting, each distributor can benefit at the expense of the SO. For example, consider a reward scheme R such that $a_R \left(1 \right) > \phi_R$. By misreporting that he sold a distribution license to some fake identity j, distributor i has to pay j's license fee ϕ_R in order to obtain a commission of $a_R \left(1 \right) > \phi_R$.

---

[8]Allowing for an endogenous price would have no effect on our results.

Comment: Purchases for own use

In our model, agents have one opportunity to purchase the good. An agent who rejects an offer to purchase the good cannot purchase it at a later stage, even if he becomes a distributor himself. It would be more realistic to assume that agents who become distributors can purchase the good for their own use at a price of $\eta_R - b_R(1)$ (i.e., to "sell" the good to themselves and obtain the commission for that sale). We have chosen to preclude this option since doing so greatly simplifies the exposition and allows us to deliver the intuition for some of the proofs in a clearer way. Our modeling choice has no effect on any of our results (proofs of our results without this simplification are available upon request).

The game. The timeline in our model is as follows. First, the SO chooses a reward scheme R. Then, the agents and the SO play the distribution game $\Gamma(R)$ that is induced by the scheme. We assume that each agent knows at what time t he enters the game and the rules of the game. To solve $\Gamma(R)$, we use subgame perfect Nash equilibrium (SPE).

We now state two observations on the SPEs of $\Gamma(R)$. The first observation follows from the fact that our setting is not stationary: as time passes, the number of agents whom a new entrant is expected to meet goes to 0.

Observation 1 *There exists a time $t^\star(R)$ such that at each time $t > t^\star(R)$, every agent who receives an offer to purchase a distribution license rejects it.*

Proof. Consider an agent j who enters the game at time t. His expected number of descendants (in G) is smaller than $\sum_{i=1}^{\infty} \frac{\delta}{t+1}$. As this expression goes to 0 when t tends to infinity, j's expected number of descendants (in G) goes to 0 as well. Incentive compatibility implies that $R(T_j) \leq (\phi_R + \eta_R)(|G_j| - 1)$. Since $c > 0$, there is a round $t^\star$ from which time onwards agents never purchase a license. ∎

When an agent is indifferent between accepting and rejecting (respectively, making and not making) an offer to purchase the good or a license, then we break this indifference and assume that he accepts it. This implies both unique and symmetric SPE behavior by the agents (the agents' decisions are independent of their labels and the labels of agents with whom they interact). We denote the SO's expected profit in the SPE of $\Gamma(R)$ by $\pi(R)$.

Observation 2 *If $q > 0$ and $\eta_R < 1$, then there exists a scheme R', identical to R except that $\eta_{R'} \in (\eta_R, 1]$, such that $\pi(R') > \pi(R)$.*

Proof. Consider the SPE that was played in $\Gamma(R)$ and fix the agents' strategies. We show that these strategies are a part of an SPE of $\Gamma(R')$. Under both schemes, for every $i \in I$, selling the good is always a best response and buying it is a best response if and only if $\omega_i = 1$. Consider an agent $i \in I$ who decides whether or not to purchase (respectively, sell) a license at time t of $\Gamma(R')$. Since the commissions in both schemes are identical and we fixed the other agents' profile of strategies, it is a best response for i to purchase (respectively, sell) a license if and only if it is a best response to do so at time t of $\Gamma(R)$. Hence, the profile of strategies that forms the SPE that was played in $\Gamma(R)$ also forms an SPE of $\Gamma(R')$. It follows that by perusing the same strategy in both games the SO can make the same transfers to the distributors, receive the same amount of license fees, and sell the same amount of the good. Since $\eta_{R'} > \eta_R$, the revenue obtained by the SO in $\Gamma(R')$ is strictly greater than the revenue obtained by the SO in $\Gamma(R)$. Hence, $\pi(R') > \pi(R)$. ∎

In light of Observation 2, we shall restrict attention to schemes R such that $\eta_R = 1$ and omit η_R from the description of the scheme.

3   Benchmark: Fully Rational Agents

3.1   The case of $q = 0$

We formally show that when $q = 0$ (i.e., there is no demand for the good), the SO cannot make a strictly positive expected profit. In this case, each reward scheme induces a bet between the SO and the participants on the number of agents who will purchase a license. The no-trade theorems (Milgrom and Stokey, 1982; Tirole, 1982) imply that if agents are willing to purchase licenses, then the SO's profit is non-positive.

Proposition 1 *Let $q = 0$. There exists no reward scheme R such that $\pi(R) > 0$.*

Proof. Consider a reward scheme R. Since $q = 0$, the game $\Gamma(R)$ induces transfers between the SO and agents who purchase a distribution license. Each of the latter agents must have an expected payoff of at least 0 in every SPE of $\Gamma(R)$. Hence, if agents purchase licenses in an SPE, then $\pi(R) < 0$. ∎

If the SO can make a strictly positive expected profit when there is no genuine product to be traded (i.e., $q = 0$), then we say that he can *sustain a pyramid scam*. Proposition 1 confirms that the SO is unable to sustain such a scam when he faces fully rational agents.

3.2 The optimal scheme when $q > 0$

In this subsection, we relax the assumption that $q = 0$ and adopt the point of view of a profit-maximizing SO. First, we examine whether the SO would benefit from using a scheme that charges entry fees or offers recruitment-based pay. The next proposition establishes that a profit-maximizing scheme need not offer recruitment-based compensation or charge entry fees.

Proposition 2 *There exists an incentive-compatible reward scheme R such that $\phi_R = 0$, $a_R(\tau) = 0$ for every $\tau \in \mathbb{N}$, and $\pi(R) \geq \pi(R')$ for every other reward scheme R'.*

Proof. The first part of the proof considers an arbitrary scheme R' and shows that if $\phi_{R'} > 0$, then there exists another scheme R such that $\phi_R = 0$ and $\pi(R) \geq \pi(R')$. Observe that R need not be incentive compatible. In the second part of the proof, we shall construct an incentive-compatible scheme $R^\star$ such that $\pi(R^\star) \geq \pi(R)$.

Consider $\Gamma(R')$, and, without loss of generality, assume that the agents are indifferent between purchasing a license and not doing so in at least one round which we shall denote by t (if there were no such round, then either the SO would be able to increase his profit by raising the entry fees, or no agent would purchase a license). Construct another scheme R such that $a_R(\tau) = \gamma a_{R'}(\tau)$ and $b_R(\tau) = \gamma b_{R'}(\tau)$ for every $\tau \in \mathbb{N}$, and $\phi_R = 0$. We can set $\gamma < 1$ such that, in round t of $\Gamma(R)$, the agents are exactly indifferent between purchasing a license and not doing so, given the strategies that were played in $\Gamma(R')$. In expectation, given these strategies, this scaling down weakly reduces the net transfers (the commissions minus the entry fees) from the SO to agents whose expected payoff is positive in the SPE of $\Gamma(R')$, while keeping the expected payoff non-negative. Since the only change made in the transition from R' to R is a scaling down of the commissions and the fees, this profile of strategies is part of the SPE of $\Gamma(R)$. This completes the first part of the proof.

Consider an arbitrary reward scheme R such that $\phi_R = 0$. By Observation 1, in $\Gamma(R)$, there is a last round in which agents purchase a license, which we shall denote by $t^\star$. Let $R^\star$ be such that for every directed weighted tree T_i,

$$R^\star(T_i) = \sum_{j \in T_i} p^{d_T(i,j)} b_R(1) w_j \tag{2}$$

Under $R^\star$, each distributor is indifferent between selling a license to agents whom he meets and not doing so. This is because, for each agent $i \in I$, independently of the number of distributors in his down-line, the expected revenue from selling (possibly

11

indirectly) a unit of the good to an agent $j \in G_i$ such that $d_G(i, j) = x$ is $p^{x-1}b_R(1)$. Since $b_R(1) = b_{R^\star}(1)$, each agent who purchases a license in round τ of $\Gamma(R^\star)$ receives, in expectation, the same transfers as he would receive in $\Gamma(R)$ following the strategy of offering a unit of the good to each agent whom he meets and not selling licenses at all. Thus, the expected transfer to an agent who purchases a license in round τ is weakly greater under R. We now use this fact to show that $R^\star$ mimics the solution to an auxiliary problem that guarantees the SO an expected profit greater than $\pi(R)$.

Consider an auxiliary problem in which the SO chooses a subset of rounds $S \subseteq \{1, ..., t^\star\}$ such that in each round $t \in S$, a distributor who meets an agent must make the latter an offer to purchase a license and the latter must accept the offer. In each round $t \notin S$, a distributor who meets an agent is not allowed to make him an offer to purchase a license. Distributors are paid according to $R^\star$ in the auxiliary problem and each agent i purchases the good if and only if $\omega_i = 1$. Since the SO can decide when agents will purchase licenses and because his expected transfers are weakly lower than his expected transfers under R, his payoff in the auxiliary problem is at least $\pi(R)$.

Under $R^\star$, the expected net profit (the revenue minus commissions paid to distributors) of the SO from the potential sale to agent j such that $d_G(SO, j) = x \geq 2$ and there are $y \in \{0, 1, ..., x - 2\}$ distributors on the directed path from him to j is

$$qp^{x-1-y}[1 - \sum_{t=1}^{y} p^{y-t}b_{R^\star}(1)] \tag{3}$$

Adding a distributor to the directed path would raise the SO's net profit if and only if $b_R(1) \leq 1 - p$. Hence, the solution to the auxiliary problem is $S = \{1, ..., t^\star\}$ if $b_R(1) \leq 1 - p$ and $S = \emptyset$ if $b_R(1) > 1 - p$. In the former case, the SPE of $R^\star$ imitates the solution as it is a best response to purchase (and sell) a license in every round $t \leq t^\star$. In the latter case, any reward scheme that incentivizes the agents not to purchase licenses imitates the solution to the auxiliary problem. ∎

At the heart of this proof is a scaling-down argument. A profit-maximizing SO is always better off if he can scale down the (strictly positive) net transfers he makes to the distributors without affecting their behavior. When a scheme charges entry fees, it is always possible for him to do so.

The incentive-compatible scheme we constructed in the proof is geometric (Emek et al., 2011). A scheme R is said to be *geometric* if there is an $s \leq 1$ such that for every $\tau \in \mathbb{N}$, $\frac{b_R(\tau+1)}{b_R(\tau)} = s$. The optimality of the geometric scheme $R^\star$ follows from the fact that each distributor is paid the exact amount that he can secure by making only direct retail sales (without selling licenses) given $b_{R^\star}(1)$. Observe that this is the

minimal compensation that is required to incentivize a distributor to recruit a new distributor and lose access to the latter's descendants in G.

We say that two schemes, R and R', are *equivalent* if $\pi(R) = \pi(R')$ and for every round τ, an agent who enters $\Gamma(R)$ at time τ makes the same expected payoff as an agent who enters $\Gamma(R')$ at time τ. The next proposition establishes that every optimal scheme is equivalent to some geometric scheme.

Proposition 3 *Suppose that R is an incentive-compatible scheme that maximizes the SO's expected profit. There exists an incentive-compatible geometric reward scheme that is equivalent to R, does not charge entry fees, and does not pay recruitment-based compensation.*

Proof. There are two cases: either $\phi_R > 0$, or $\phi_R = 0$. First, suppose that $\phi_R = 0$. Since R is optimal, it must mimic the solution to the auxiliary problem presented in the proof of Proposition 2. Hence, either R is degenerate (i.e., agents do not purchase a license on the equilibrium path of its induced game) or agents purchase a license in rounds $1, ..., t^\star$ of $\Gamma(R)$ and each of them is paid the minimal amount possible given $b_R(1)$. In the latter case, R is equivalent to the geometric scheme given in (2). In the former case, it is equivalent to a geometric scheme that pays $b_{R^\star} = 0$.

To complete the proof, suppose that $\phi_R > 0$ and consider the equilibrium path of $\Gamma(R)$. There are two cases: agents purchase a license in exactly one round t, or they purchase a license in at least two rounds. In the former case, it is easy to construct a geometric scheme that makes agents indifferent between purchasing a license in round t and not doing so. Recall that the SO may refrain from making offers in every round $t' < t$. Hence, either such a geometric scheme outperforms R, or it guarantees the same expected transfers to a distributor who purchases a license in round t.

Let us show that R cannot maximize the SO's expected profit if there are two rounds in which agents purchase a license on the equilibrium path of $\Gamma(R)$. Recall the scaling-down argument from step 1 of the proof of Proposition 2. The SO is strictly better off when he scales down the commissions and the entry fees if there exists one agent who makes a strictly positive expected payoff. By our assumption, under R, agents purchase a license on the equilibrium path in two rounds, t_1 and t_2, where t_2 denotes the last round in which they purchase a license. If an agent who purchases a license in t_2 makes an expected profit of 0 from selling the good (recall that he does not sell licenses on the equilibrium path), then an agent who purchases a license in round t_1 can secure a strictly positive payoff by following the strategy of selling the good to each agent whom he meets and not selling licenses at all. ■

Proposition 3 is useful since it allows us to restrict attention to geometric reward schemes. We now study the income distribution that is induced by the optimal geometric reward scheme.

The income distribution

Denote by z_t the expected number of agents whom an agent who purchases a license in round t meets if he does not sell any distribution license. Observe that the last agent who purchases a license in the SPE of $\Gamma(R^\star)$ (in round $t^\star$) makes an expected payoff of 0 (otherwise, it would be possible to charge entry fees and increase the SO's profit). Thus, the expected payments from the SO to that distributor sum up to $c = qb_{R^\star}(1) z_{t^\star}$. That is, $b_{R^\star}(1) = \frac{c}{z_{t^\star}q}$. The distributor who purchases a license in round $t < t^\star$ earns an expected net payoff of $c\left(\frac{z_t}{z_{t^\star}} - 1\right)$. Since

$$z_t = \left(\frac{\delta}{t+1} + \frac{\delta^2}{t+2} + ...\right) + \frac{1}{t+1}\left(\frac{\delta^2 p}{t+2} + \frac{\delta^3 p}{t+3} + ...\right) + ... \tag{4}$$

is an infinitely countable sum of functions that are decreasing and convex in t, it follows that the expected net income is decreasing and convex with respect to the time at which agents purchase a license and become distributors.

Multiple layers of intermediation

The SO has to choose how many distributors to use (i.e., he has to choose $t^\star$, which in turn pins down $b_{R^\star}(1)$). The marginal benefit of using the k-th distributor (i.e., the addition to the SO's revenue when he extends S from $\{1, ..., k-1\}$ to $\{1, ..., k\}$) is $(1-p)z_k$. Observe that z_k is increasing in δ and, since the harmonic sum diverges, tends to infinity at the $\delta = 1$ limit. The marginal cost of using the *second* distributor is $\frac{cz_1}{z_2}$. For large values of δ, $\frac{z_1}{z_2} \leq 2$. Therefore, there exists a number $\delta^\star$ such that for every $\delta > \delta^\star$, in a game that is induced by an incentive-compatible reward scheme that maximizes the SO's profit, agents purchase a license in rounds $t = 1$ and $t = 2$ (at least). The next proposition complements Proposition 2 by showing that if $\delta > \delta^\star$, then incentive-compatible schemes that charge entry fees or pay recruitment-based compensation do not maximize the SO's profit.

Proposition 4 *Let $p < 1$, $\delta > \delta^\star$, and consider an incentive-compatible scheme R. If R maximizes the SO's expected profit, then $\phi_R = 0$ and $a_R(1) = a_R(2) = ... = 0$.*

Proof. We shall prove this claim by showing that an incentive-compatible scheme R' such that $\phi_{R'} > 0$ or $a_{R'}(\tau) > 0$ for some $\tau \in \mathbb{N}$ cannot maximize the SO's expected profit. Let R' be an incentive-compatible reward scheme such that $a_{R'}(\tau) > 0$ for some $\tau \in \mathbb{N}$. By incentive compatibility, $\phi_{R'} > 0$. Since $\delta > \delta^{\star}$, the optimality of R' requires that on the equilibrium path of $\Gamma(R')$ agents purchase a license in a set of rounds $\{1, ..., k\}$, where $k \geq 2$. In an SPE, an agent who purchases a license in round k makes an expected payoff of at least 0 and agents who purchase a license in round $t < k$ make a strictly positive expected payoff since $z_t > z_k$. Recall the scaling-down argument from step 1 of the proof of Proposition 2. If some agents make a strictly positive expected payoff, then the SO can strictly increase his expected profit by scaling down the commissions and the entry fees (setting $\phi = 0$ and multiplying the commissions by a constant such that in round k agents are indifferent between purchasing a license and not doing so). ∎

Observe that the scaled-down scheme we constructed in the proof need not be incentive compatible. However, by Proposition 2, there exists an incentive-compatible geometric scheme that guarantees the SO an expected profit at least as great as the expected profit induced by the scaled-down scheme.

Propositions 2 and 4 establish that, in our model, for large values of δ, in an optimal scheme, only retail sales (direct and indirect ones) are compensated when $q > 0$. In other words, an incentive-compatible scheme that charges entry fees or pays recruitment commissions cannot be optimal when $q > 0$. In the next section, we shall depart from full rationality and show that entry fees and recruitment-based pay play a crucial role when agents are partially sophisticated.

4 Partially Sophisticated Agents and Pyramid Scams

Our objective in this section is to understand the mechanisms that underlie pyramid scams and the conditions that enable such scams to operate. In a real-world pyramid scam, the only "product" that is being traded is the right to join the pyramid. To reflect that, we shall assume through most of the analysis in this section that $q = 0$ (i.e., there is no demand for the good). In Proposition 1, we established that when the agents are fully rational and $q = 0$, the SO cannot make a strictly positive expected profit. A possible interpretation of that result is that fully rational agents cannot be lured into taking part in a pyramid scam. However, in settings such as ours, where agents have to consider and forecast multiple variables in order to reach a sensible decision, the full rationality assumption might be too extreme. In this section we

15

depart from the classic rational expectations model.

In the context of our model, the literature offers two potential trajectories to depart from full rationality: coarsening the agents' strategic reasoning and relaxing the common prior assumption. In our setting, agents have to forecast multiple decisions taken by their successors. These decisions depend, among other factors, on the subtleties of the reward scheme and on the way other agents grasp those subtleties. In such complicated problems, it makes sense to use simplified models of other players' behavior and information. Ideas in this spirit have been formalized by Jehiel (2005), Eyster and Rabin (2005), and Jehiel and Koessler (2008). We follow this approach and use analogy-based expectations equilibrium (Jehiel, 2005) to solve our model as it is tailored to sequential-move games such as ours. In Section 5, we discuss the alternative approach (i.e., relaxing the common prior assumption) and its implications for our results.

One interpretation of analogy-based expectations in our model is that in order to simplify their decision problems, the players create simplified representations of their opponents' behavior and reason entirely in terms of these representations. Alternatively, the formation of analogy-based expectations can be interpreted as *learning from coarse feedback* about past play (in similar schemes). We shall elaborate on the latter interpretation after we formally present the solution concept.

4.1 Coarse reasoning

In this subsection we apply analogy-based expectations equilibrium (ABEE) to our model. There are four types of decisions in $\Gamma(R)$: purchasing a license, selling a license, purchasing a unit of the good, and selling a unit of the good. Observe that there are more decision nodes in which an agent makes the second (fourth) type of decision than there are decision nodes in which an agent makes the first (third) type of decision.

Formally, denote by M_1 the set of histories after which players decide whether to purchase a distribution license or not, by M_2 the set of histories after which players decide whether to buy a unit of the good or not, by M_3 the set of histories after which players decide whether to offer a distribution license or not, and by M_4 the set of histories after which players decide whether to offer a unit of the good or not. We refer to M_1, M_2, M_3, and M_4 as the players' analogy classes.

For each agent $i \in I$, we denote i's strategy by σ_i, and his (possibly mixed) action in each node $m \in M_1 \cup M_2 \cup M_3 \cup M_4$ by $\sigma_i(m) \in [0, 1]$. The decision to accept/make

(respectively, reject/not make) an offer to purchase a license or the good in node m is associated with $\sigma_i(m) = 1$ (respectively, $\sigma_i(m) = 0$). Denote by σ_{SO} the SO's strategy, let $\sigma := (\sigma_i)_{i \in I \cup \{SO\}}$, and denote the probability of reaching node m given σ by $r_\sigma(m)$. Observe that the SO's strategy affects[9] $r_\sigma(m)$. For each node $m \in M_1 \cup M_2 \cup M_3 \cup M_4$, denote by $\sigma(m)$ the expectation over the agents' actions in m that is induced by the profile σ.

For each $k \in \{1,2,3,4\}$ and $i \in I$, let $\beta_i(M_k)$ denote agent i's *analogy-based expectations* over his opponents' behavior in M_k. Let $\beta_i := (\beta_i(M_k))_{k \in \{1,2,3,4\}}$ and $\beta := (\beta_i)_{i \in I}$.

Definition 1 *Player i's analogy-based expectations β_i are said to be consistent with σ if for every $k \in \{1,2,3,4\}$, $\beta_i(M_k) = \frac{\sum_{m \in M_k} r_\sigma(m)\sigma(m)}{\sum_{m \in M} r_\sigma(m)}$ whenever $r_\sigma(m) > 0$ for some $m \in M_k$.*

In words, the consistency of the players' analogy-based expectations implies that these expectations *match the average behavior* in the different analogy classes under σ.

Note that consistency implies that the agents assign lower weights to histories that are less likely to be reached. To illustrate this, let $\delta = 0.5$, $p = 0$, and consider a profile σ in which agents always purchase (respectively, never purchase) a distribution license in round 1 (respectively, round $t > 1$). Further, suppose that distributors (and the SO) who meet agents at time $t \in \{1,2,3\}$ always make them an offer to purchase a license. When a distributor (or the SO) meets an agent at $t > 3$, the former does not make an offer to the latter. Under this profile, at time $t \in \{1,2\}$, agents receive an offer to purchase a license with probability 1. An agent who enters the game at time $t = 3$ will receive such an offer with probability $\frac{2}{3}$ since there is a probability of $\frac{1}{3}$ that he will meet the agent who entered the game at round 2 who is not a distributor. Recall that time t is reached with probability δ^t. Consistency implies that $\beta_i(M_1) = \frac{\delta}{\delta + \delta^2 + \frac{2}{3}\delta^3} = \frac{3}{5}$. Observe that β depends on the SO's behavior. When the SO makes offers in rounds 2 and 3, more agents are likely to receive offers and reject those offers.

Definition 2 *The strategy σ_i is a best response to β_i if it best responds to the belief that for each $k \in \{1,2,3,4\}$ and every node $m \in M_k$, each agent $j \in I/\{i\}$ plays $\sigma_j(m) = \beta_i(M_k)$.*

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

---

[9]Whether the agents' understanding of the SO's strategy is coarse or not has no effect on the analysis.

Observe that, given a scheme R, there might be more than one ABEE in $\Gamma(R)$. Let us redefine $\pi(R)$ to be the *maximal expected profit* that the SO can make in an ABEE of $\Gamma(R)$.

We can interpret agents who base their decisions on their analogy-based expectations as agents who *learn from coarse feedback* on past play in *similar schemes*. For instance, they might receive data about the proportion of offers to purchase a license that was accepted in these schemes (the SO might provide them with this information himself). The information received by our agents does not allow them to perceive subtleties in the data such as the connection between purchasing decisions and time. In fact, they might not even be aware of these complex statistical relations. Our agents best respond to the coarse data that they do receive. The consistency requirement implies that the feedback that the agents do obtain is correct on average. This may reinforce their belief in the simplified model that they have in mind.

4.2   The case of $q = 0$

Our objective in this subsection is to understand the properties of the reward schemes that underlie pyramid scams and the conditions that allow such scams to work. In order to model the fact that in a real-world pyramid scam there is no genuine product and therefore participants can trade only the right to join the pyramid, we assume throughout this subsection that $q = 0$ (this implies that compensation that is based on retail sales has no effect on the agents' decisions). As in Section 3.1, if there exists a reward scheme R such that $\pi(R) > 0$, then we say that the SO can sustain a pyramid scam.

First, we investigate the relation between the number of *layers* on which the distributors' compensation is based and the possibility that the SO makes a strictly positive expected profit. We show that a reward scheme that underlies a pyramid scam must offer compensation that is based on more than one level of the agents' down-lines.

Definition 4 *A reward scheme R is said to be a one-level scheme if $a_R(\tau) = 0$ and $b_R(\tau) = 0$ for every $\tau > 1$.*

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

Proof. First, note that Observation 1 holds. Therefore, if agents purchase licenses in an ABEE of $\Gamma(R)$, then there exists a last round in which they do so. Second, if it is a best response for an agent to purchase a license in round t, then it is a best response

18

for him to do so in each round $t' < t$. Let us assume by way of contradiction that there exists in $\Gamma(R)$ an ABEE in which agents purchase licenses and denote by k the last round in which an agent purchases a license in this ABEE.

Denote by s_k the expected number of offers (to purchase a license) made in this ABEE in rounds $1, ..., k$ by the SO and the distributors. On average, s_k agents hold a distribution license after k rounds. In the rounds that follow round k, agents reject every offer to purchase a license. Let $\hat{v}_k$ denote the number of offers that the agent who purchases a license in round k expects to make in those rounds (according to his analogy-based expectations). Let v_k denote the expected number of offers this agent makes. Note that $\hat{v}_k \leq v_k$ since the agents overestimate the probability that the offers that they make will be accepted in rounds $t > k$ (when an offer is rejected by agent j, if $p > 0$, the agent who made the offer may meet the members of G_j and make them some additional offers). Agents who purchased a license before round k make, on average, at least v_k offers after round k (agents who decided not to make offers in previous rounds are expected to meet more new entrants in rounds $t > k$ and therefore are expected to make more offers). Since offers in rounds $t > k$ are always rejected,

$$\beta_i(M_1) \leq \frac{s_k}{s_k + s_k v_k} \tag{5}$$

for each $i \in I$. The incentive compatibility of R implies that $a_R(1) \leq \phi_R$. As a result, the agent who enters the game in round k believes that he will receive commissions of

$$a_R(1) \beta_i(M_1) \hat{v}_k \leq \frac{a_R(1) s_k \hat{v}_k}{s_k + s_k v_k} \leq \frac{\phi_R v_k}{1 + v_k} < \phi_R \tag{6}$$

Hence, it is not a best response to purchase a license in round k. This stands in contradiction to the assumption that agents purchase a license in round k in this ABEE. ∎

Since $q = 0$, a distribution license is essentially a bet on the number of agents who will purchase a license in the future. Even though the agents' analogy-based expectations are typically overoptimistic, the SO cannot lure the agents into taking such a bet by using incentive-compatible one-level schemes. That is, it takes more than a simple scheme to take advantage of our agents' coarse reasoning.

To obtain an intuition for this result, consider an agent i who receives an offer to purchase a license in round k (which is the last round in which offers are accepted in the assumed ABEE). Given an incentive-compatible one-level scheme, i will accept the offer only if the number of licenses he expects to sell, $\hat{v}_k \beta(M_1)$, is greater than 1.

The key point is that every offer that contributes to $\hat{v}_k$ is made after round k, and is, therefore, rejected. This rejection leads to a reduction in $\beta(M_1)$, which is the ratio between the number of accepted offers and the total number of offers made (we can interpret this reduction as a negative "feedback" about the likelihood that other agents' purchase a license when they have the opportunity to do so). Thus, even though $\hat{v}_k$ can be large, $\hat{v}_k\beta(M_1)$ cannot exceed 1 and hence i prefers not to purchase a license in round k.

Can the SO sustain a pyramid scam when he is allowed to use schemes that reward multiple layers of "intermediation"? The answer to this question depends on p and the expected number of entrants. In particular, if δ is sufficiently large and p is sufficiently small, then a pyramid scam is sustainable.

Proposition 6 *Fix $q = 0$. There exists a number $\delta^\star < 1$ such that for every $\delta > \delta^\star$ there exists a number $p^\star(\delta) < 1$ such that:*

1. *If $p \leq p^\star(\delta)$, then there exists an incentive-compatible scheme R such that $\pi(R) > 0$.*

2. *If $p > p^\star(\delta)$, then there exists no incentive-compatible scheme R such that $\pi(R) > 0$.*

Proof. First, we establish that in the extreme case of $p = 1$, there exists no scheme R such that $\pi(R) > 0$. Assume to the contrary that there exists a scheme R such that $\pi(R) > 0$. Then, there is an ABEE of $\Gamma(R)$ in which there is a last round in which agents accept offers to purchase a license; denote this round by k. By (6), the analogy-based expectations of agents who purchase a license in round k are that they will have less than one distributor in the first level of their down-line. According to these expectations, their successors (i.e., agents who purchase a license after round k) will have less than one distributor in the first level of their down-lines. Therefore, it is a best response for each distributor to offer a license to every agent whom he meets from round k onwards.

Let s_k be the expected number of offers made in the first k rounds in the assumed ABEE. The expected number of agents in the subtree G_i rooted at the agent i who purchases a license in round k is denoted by E_k. Since $p = 1$ and from round $k + 1$ onwards agents reject every offer that they receive, the expected number of offers each distributor makes from round $k + 1$ onwards is at least E_k. Hence, $\beta(M_1) \leq \frac{1}{1+E_k}$. The expected number of distributors in the *down-line* of a distributor who purchases

a license in round k cannot exceed

$$\frac{E_k}{1 + E_k} < 1 \tag{7}$$

By incentive compatibility, $R(T_i)$ cannot exceed $(|T_i| - 1) \phi_R$ for every directed weighted tree T_i. Therefore, the expected payoff received by the distributor who purchases a license in round k is at most

$$\frac{E_k \phi_R}{1 + E_k} - \phi_R < 0 \tag{8}$$

This stands in contradiction to the assumption that there exists an ABEE in $\Gamma(R)$ in which agents purchase a license in round k.

The next step of the proof is to analyze the other extreme case. That is, the case that $p = 0$. We show that there exists a number $\delta^\star$ such that for $\delta > \delta^\star$, there exists an incentive-compatible scheme R such that $\pi(R) > 0$. We shall prove this by constructing a scheme such that in its induced game there exists an ABEE that guarantees the SO a strictly positive expected profit. Consider a scheme R such that for each directed weighted tree T_i,

$$R(T_i) = x(|T_i| - 1) - \phi_R \tag{9}$$

Let $x \leq \phi_R$. It follows that R is incentive compatible. Consider a profile of strategies σ such that:

- An agent who receives an offer to purchase a license in round 1 accepts it.

- Agents reject offers in every round $t > 1$.

- The SO offers a distribution license only in round 1.

- Distributors offer distribution licenses in every round $t > 1$.

This profile guarantees the SO an expected profit of ϕ_R. We now show that if δ is sufficiently large, then σ is an ABEE.

Consider a distributor i who meets a new entrant j. Given R, i's objective is to maximize the size of his down-line. Hence, distributor i will make an offer to agent j to purchase a license regardless of the time at which they meet. Note that in every round $t > 1$ the SO is indifferent between making an offer and not doing so. It is left to show that purchasing a license is a best response in round 1 and is not a best

response in every round $t > 1$. Since the expected payment to an agent who purchases a license in round 1 is strictly greater than the expected payoff obtained by an agent who purchases a license in round 2, it is sufficient to show that there exists $x < \phi_R$ such that it is a best response for an agent to purchase a license in round 1.

Consider an agent j who purchases a license in round 1. Since he always makes offers to other agents, if the game lasts up to period $t = n$, then, in expectation, he makes $\sum_{i=2}^{n} \frac{1}{i}$ offers. Consistency implies that agent j believes that he will have, on average,

$$\frac{\delta \sum_{i=1}^{\infty} \frac{\delta^i}{1+i}}{\delta + \delta \sum_{i=1}^{\infty} \frac{\delta^i}{1+i}} < 1 \tag{10}$$

distributors in the first level of his down-line. Since the harmonic sum diverges, the LHS of (10) tends to 1 at the $\delta = 1$ limit. Consistency also implies that $\beta_j(M_3) = 1$. Thus, according to his analogy-based expectations, j believes that he will have

$$L_k = \frac{\sum_{j_1=2}^{\infty} \sum_{j_2=j_1+1}^{\infty} \cdots \sum_{j_k=j_{k-1}+1}^{\infty} \frac{\delta^{j_k}}{j_1 j_2 \ldots j_k}}{\left(1 + \sum_{i=1}^{\infty} \frac{\delta^i}{1+i}\right)^k} \tag{11}$$

distributors in the k-th level of his down-line. As δ approaches 1, L_k tends to $\frac{1}{k!}$. Hence, there exists a number $\delta^\star$ such that $\sum_{k=1}^{\infty} L_k > 1$ for every $\delta > \delta^\star$. For every such δ, it is possible to set x below ϕ_R in a way that would make the first entrant indifferent between purchasing a license and not doing so and would make the second entrant strictly prefer not to purchase a license to doing so. Observe that this result is true for every $c > 0$ since we can always set x below ϕ_R and multiply both by a constant $\gamma > 1$.

To complete the proof, fix $\delta \geq \delta^\star$, $p = p^\star$, and suppose that there exists an incentive-compatible scheme R such that $\pi(R) > 0$. We show that for every $p < p^\star$, there exists an incentive-compatible scheme R' such that $\pi(R') > 0$. Denote by k the last round of $\Gamma(R)$ in which a license is purchased in an ABEE in which the SO's profit is strictly positive. Denote the agents' analogy-based expectations in this ABEE by $\beta_1' = \beta(M_1)$ and $\beta_3' = \beta(M_3)$. Consider a reward scheme R' that pays $\phi_R(|T_i| - 1) - \phi_R$ for every tree T_i. Observe that R' is incentive compatible. Fixing $\beta(M_1) = \beta_1'$ and $\beta(M_3) = \beta_3'$, the k-th entrant's (the agent who purchases a license in round k) analogy-based expectations are that his expected payoff is at least as high as it is in the ABEE of $\Gamma(R)$ that we considered. Raising $\beta(M_3)$ above β_3' would only increase the k-

th entrant's payoff (given R' and $\beta(M_1) = \beta'_1$) since it would increase the expected number of agents in the k-th entrant's down-line, which is his objective under R'.

Consider the game $\Gamma(R')$ and the following profile of strategies σ. The SO offers a license in a subset of rounds K, where $k \in K$ and $k \geq t$ for each $t \in K$. Distributors always make offers to agents whom they meet. Agents accept offers to purchase a license in rounds $1, ..., k$ and reject such offers from round $k+1$ onwards. We shall argue that σ and a consistent profile of analogy-based expectations β form an ABEE in which the SO's expected profit is strictly positive. This will allow us to focus on the scheme R' when we prove our claim.

First, let us consider the SO. Making an offer in the k-th round is clearly part of a best response as the SO obtains entry fees and does not pay commissions to the k-th entrant. Not making offers from round $k+1$ onwards is a best response since the agents reject such offers. If the SO makes an offer only to the k-th entrant, then his expected profit is strictly positive. Hence, the SO's expected profit in this ABEE is strictly positive.

The analogy-based expectations $\beta(M_3) = 1$ are consistent with σ. Further, since the distributors always make offers to agents whom they meet under σ, a consistent $\beta(M_1)$ is independent of the expected number of offers that the SO makes in rounds $1, ..., k$ as this number cancels out in $\beta(M_1)$ (see, e.g., the RHS of (5)). Further, $\beta(M_1) \geq \beta'_1$ as the expected number of offers made after round k (per distributor) under R' is weakly greater than the respective number of offers per distributor under R (since under R there might be distributors who did not make an offer to an agent whom they met before round k and therefore after round k they are expected to meet a larger number of agents, which raises the proportion of rejections and lowers $\beta(M_1)$). Since $\beta(M_3) \geq \beta'_3$ and $\beta(M_1) \geq \beta'_1$, it is a best response to purchase a license in every round $t \leq k$ of $\Gamma(R')$. To make it a best response not to purchase a license from round $k+1$ onwards, we can multiply the commissions $a_{R'}(1), ..., a_{R'}(n), ...$ by a constant $\gamma < 1$ without changing the equilibrium behavior. Observe that offering a license is always a best response given R'.

To conclude, if there exists a scheme R such that in its induced game there exists an ABEE in which the SO's expected profit is strictly positive, then in the game induced by R' there exists an ABEE in which $\beta(M_3) = 1$ and the k-th entrant purchases a license on the equilibrium path. We now focus on ABEEs of this kind, and show that such an ABEE (with a corresponding incentive-compatible scheme R') exists for every $p < p^\star$. Our restriction to this type of ABEEs allows us to fix $\beta(M_3) = 1$ and $\beta(M_1)$, and focus on the k-th entrant's payoff.

Denote the k-th agent to enter the game by i, and the expected number of agents (conditional on reaching round k) in the j-th level of the tree G_i by l_j. For example, $l_1 = \sum_{j=1}^{\infty} \frac{\delta^j}{k+j}$ and $l_2 = \sum_{i'=1}^{\infty} \sum_{j'=i'+1}^{\infty} \frac{\delta^{j'}}{i'j'}$. Observe that $\beta(M_1) = \frac{1}{1+v(p)}$, where

$$v(p) := v = l_1 + pl_2 + p^2 l_3 + p^3 l_4 + ... \tag{12}$$

The expected number of distributors in i's down-line according to his analogy-based expectations is

$$\frac{l_1}{(1+v)} + \frac{l_2(1+pv)}{(1+v)^2} + \frac{l_3(1+pv)^2}{(1+v)^3} + \frac{l_4(1+pv)^3}{(1+v)^4} + ... \tag{13}$$

As long as (13) is strictly greater than 1, we can set $\phi_{R'}$ and $a_{R'}(1) = a_{R'}(2) = ... \leq \phi_{R'}$ such that σ and β form an ABEE as described above. If (13) is less than or equal to 1, we cannot do so (the fact that $c > 0$ is the reason for the strict inequality). In the appendix, we show that the derivative of (13) with respect to p is strictly negative. Hence, if for $\delta \geq \delta^\star$ and $p = p^\star$ there exists a scheme R such that $\pi(R) > 0$, then for δ and every $p < p^\star$, there exists a scheme $\hat{R}$ such that $\pi(\hat{R}) > 0$. ∎

The fact that the SO is allowed to condition his pay on multiple layers of agents' down-lines allows him to sustain a pyramid scam when δ is large and p is small. We shall now provide some intuition regarding the roles that p, δ, and multilevel compensation play in this result.

Let us consider the negative effect of p on the SO's ability to sustain a pyramid scam. In an ABEE, the agents accept (respectively, reject) offers to purchase a license in every round $t \leq k$ (respectively, $t > k$). The higher p is, the more offers are made after round k since the entrants in those rounds are more likely to run into distributors. This reduces the proportion of accepted offers (i.e., the agents' analogy-based expectations $\beta(M_1)$). Observe that a higher p may also increase the number of offers made before round k and that such offers are accepted. However, this effect does not lead to an increase in the agents' analogy-based expectations. The reason for this is that every offer that is accepted before round k results in an additional distributor who makes additional offers after round k, and these additional offers are rejected and completely cancel the positive effect that the accepted offer has on $\beta(M_1)$. In conclusion, p affects the SO's ability to sustain a pyramid scam through its negative effect on the agents' analogy-based expectations that future entrants will purchase a license.

We can interpret p as people's tendency to mention a good or a service to their friends when they have no financial incentive to do so. Under this interpretation, it is

24

natural to connect useless goods and services with low values of p that allow the SO to sustain a pyramid scam.

The role multilevel compensation plays in this result is closely related to incentive compatibility. Recall that incentive compatibility implies that the commissions cannot exceed the entry fees. In a one-level scheme, this means that a necessary condition for an agent to find it beneficial to purchase a license is that he expects to recruit at least one other agent. Proposition 5 showed that, in any assumed ABEE, this condition is not met by the last agent who is supposed to purchase a license. Incentive compatibility imposes a more lenient constraint on reward schemes that offer multilevel compensation. Given such a scheme, an agent may find it beneficial to purchase a license only if he expects to have at least one more agent in his down-line.

How many layers of compensation are required to sustain a pyramid? It turns out that if δ is sufficiently large and p is sufficiently small, then two layers of compensation are sufficient to sustain a pyramid scam. The proof of the following result is similar to that of Proposition 6, and therefore we omit it.

Proposition 7 *Fix $q = 0$. There exists a number $\delta^{\star\star} > \delta^{\star}$ such that for every $\delta > \delta^{\star\star}$ there exists a number $p^{\star\star}(\delta) < 1$ such that if $p \leq p^{\star\star}(\delta)$, then there exists an incentive-compatible scheme R such that $\pi(R) > 0$ and $a_R(\tau) = 0$ for every $\tau > 2$.*

How large is $\delta^{\star}$? For $p = 0$, it is easy to answer this question by substituting values into (10) and (11). It turns out that a value of $\delta = \frac{31}{32}$, which reflects an average population of 32 agents, is sufficient for a pyramid scam to be sustainable.

4.3 Sophistication and pyramids

In the previous section we found that it is possible to sustain a pyramid scam when agents use simplified representations of other agents' strategies. A natural question to ask is whether agents with a higher degree of sophistication are more or less vulnerable to pyramid scams.

To answer the above question, we develop a cognitive hierarchy model in the spirit of Camerer, Ho, and Chong (2004), which is the standard framework for analyzing strategic interaction among agents with different degrees of sophistication. The conventional cognitive hierarchy model is built on a naive type (henceforth, L_0) who plays some naive strategy that is exogenously given (e.g., the strategy of uniformly randomizing among actions). Types are ordered according to their level of sophistication, and

type L_k, $k \geq 1$, best responds to the belief that all other agents are of types lower than k. In the typical level-k thinking model, L_0's decision rule pins down the model.

The naive type in our model *best responds to a simplified representation of his opponents' behavior.* This form of naive behavior is consistent with the model we presented in the previous section. In fact, the agents whose behavior we analyzed in the previous section correspond to type L_0 in the present section. As in the previous section, we can interpret the naive type's behavior as best responding to coarse data that is supplied by the SO (e.g., the ratio between the number of offers that were accepted and the number of offers that were rejected in similar schemes in the past).

The innovation in our cognitive hierarchy model is that the *naive type's behavior is determined in equilibrium.* It is affected by the behavior of higher types and in turn affects these higher types' behavior. Our extension of our behavioral model is in the spirit of Eyster and Rabin's (2010) best response trailing naive inference (BRTNI) play. Unlike in our model, in a BRTNI play, the naive type's behavior is exogenous as he best responds to his own private information while ignoring the informational content in other players' actions.

We now define the equilibrium concept formally. Let F be a true distribution of types. We assume that F is known to the SO. Denote the strategy of *type L_k* by σ_{L_k} and let σ denote a profile of strategies played by the SO and all of the types. We keep the definitions of analogy classes, analogy-based expectations, and consistency as in the previous section.

Definition 5 *The pair* (σ, β) *constitutes an equilibrium of* $\Gamma(R)$ *if the following conditions are met:*

- *The analogy-based expectations* β *are consistent with* σ.

- *The strategy* σ_{L_0} *is a best response to* β.

- *For every* $k > 0$, *the strategy* σ_{L_k} *is a best response to a belief that other types are distributed uniformly on* $\{0, ..., k-1\}$ *and each type L_κ plays* σ_{L_κ}.

The uniformity assumption in the third requirement of Definition 5 makes the exposition simpler. It can be replaced with much weaker assumptions without any effect on the following result. As in the previous section, there may exist multiple equilibria. We denote by $\pi(R)$ the highest equilibrium payoff that the SO can obtain in $\Gamma(R)$.

Proposition 8 *Fix* $q = 0$ *and suppose that F does not contain type L_0-agents. There exists a number* $\delta^\star$ *such that for every* $\delta > \delta^\star$, *there exists an incentive-compatible one-level scheme R such that* $\pi(R)$.

26

Proof. Consider a one-level scheme R that pays $a_R(1) \le \phi_R$. First, note that $\beta(M_3)$ has no effect on the agents' decisions since R is a one-level scheme. Second, observe that R is incentive compatible. Third, it is easy to see that for any profile σ in which some agents purchase a license, the SO's expected profit is strictly positive. We now construct an equilibrium in which agents purchase licenses.

Set σ such that agents of type L_0 purchase a license in rounds $1, ..., s_0$ and do not purchase a license in every round $t > s_0 + 1$. Since L_1 best responds to L_0, if s_0 is sufficiently late and δ is sufficiently large, then for L_1 it is a best response to purchase a license in rounds $1, ..., s_1$ and not to do so in every round $t > s_1$. Note that s_1 depends on s_0. We can recursively construct such thresholds for each type $L_2, ..., L_k$ such that $s_0 > s_1 \ge s_2 \ge s_3 \ge ... \ge s_k$. It is left to show that, for large values of δ, σ_{L_0} is a best response to the analogy-based expectations that are consistent with σ. It is sufficient to show that it is a best response for L_0 to purchase a license in round s_0 and to reject one in round $s_0 + 1$.

In round s_0, type L_0 agents care only about the number of rounds left until the end of the game and $\beta(M_1)$. All else equal, the greater δ and $\beta(M_1)$ are, the greater is L_0's expected payoff conditional on purchasing a license in round s_0. We now show that when δ is increased, it is possible to adjust σ such that $\beta(M_1)$ is kept constant.

Observe that in every round $t > s_0$, each type L_k, where $k > 0$, is indifferent between making an offer or not when he meets an agent. Suppose that each agent of these types offers a license in those rounds with probability γ. If $\gamma = 0$, then in every round $t > s_0$ no offers are made. Thus, the expected number of rounds after s_0 has no effect on $\beta(M_1)$. However, an increase in δ implies an increase in L_0's potential clientele. For instance, if the game lasts for $n > s_0$ periods, then L_0's analogy-based expected payoff is $\beta(M_1)(H_n - H_{s_0}) a_R(1) - \phi_R$, where H_n is the n-th harmonic number. Since the harmonic sum diverges at the $n = \infty$ limit, there exists a value $\delta^\star$ such that for every $\delta > \delta^\star$, it is a best response for type L_0 agents to accept offers in round s_0. We still have to show that for values of δ greater than $\delta^\star$ it is a best response for L_0 agents to reject an offer in round $s_0 + 1$.

Observe that $\beta(M_1)$ continuously decreases in γ. By (6), for $\gamma = 1$, the analogy-based expected number of distributors in the first level of the down-line of an L_0 agent who purchases a license in round $s_0 + 1$ is strictly lower than 1. By construction, $a_R(1) \le \phi_R$. Hence, for every $\delta > \delta^\star$, we can find a $\gamma(\delta) \in (0,1)$ such that it is a best response for L_0 agents to purchase a license in round s_0 and not to do so in round $s_0 + 1$. ∎

When facing a large population of agents whose degree of sophistication is strictly

greater than 0, the SO can sustain a pyramid scam. Moreover, he can do so using a one-level scheme. Further, this result is independent of p and the agents' level of sophistication (as long as the latter level is finite). Thus, even in instances in which a population of naive agents is not vulnerable to pyramid scams (i.e., when p is high or when the rewards offered in the underlying scheme are conditioned on one level of the participants' down-lines), a large population of agents whose degree of sophistication is strictly greater than 0 is vulnerable to such scams.

The intuition for this result is as follows. In the cognitive hierarchy setting, agents purchase a license since they think that they will be able to sell licenses to agents who are less sophisticated than themselves. At the bottom of this *imagined* hierarchy, there are naive agents who decide whether or not to purchase a license based on δ and $\beta(M_1)$. Unlike the model in the previous section, here the fact that the agents would reject offers from round $s_0 + 1$ onwards does not mean that $\beta(M_1)$ is decreasing in δ. This results from the fact that in the cognitive hierarchy setting, when F does not contain naive agents, offers are not made after round s_0, and, therefore, do not affect the agents' analogy-based expectations. Thus, when δ is increased, the analogy-based expected number of licenses that the imagined naive type sells increases and this incentivizes agents who are more sophisticated to purchase a license in order to take advantage of L_0's mistakes.

5 Extensions

5.1 A mechanism-design approach

Our definition of a reward scheme is with loss of generality. More generally, one could allow a reward scheme R to be any function that maps directed weighted trees to real rewards. The schemes we use have realistic motives and they allow us to focus on the multilevel nature of the problem and the commissions structure. We now discuss the changes that would result in our analysis from applying a more general approach.

It is easy to see that the possibility results (in Propositions 6–8) hold when we take a general approach. The negative result of Proposition 1 is a direct result of the no-trade theorems (Milgrom and Stokey, 1982; Tirole, 1982) and does not depend on our restrictions. It is also possible to generalize the negative result of Proposition 5 in the case of $p = 0$, which is the least favorable case for this result. When $p > 0$, our setting's tractability is significantly reduced when we allow for a general class of reward schemes.

Let us consider the benchmark results of the fully rational case (Propositions 2–4). The geometric scheme is not necessarily optimal in the general case. However, it maximizes the SO's expected profit among a larger class of schemes than the one we considered: the class of schemes that offer a proportional commission per retail sale.

Definition 6 *A reward scheme R is said to have a direct selling mode if there exist two constants $b > 0$ and $\phi \geq 0$ such that whenever $T_i = \{i\}$, then $R(T_i) = bw_i - \phi$.*

Under a scheme that offers a "direct selling" mode, distributors who focus solely on retail sales and do not recruit other distributors to their down-line obtain a fixed commission per each retail sale they make. Among the schemes that offer such a mode, the geometric scheme $R^\star$ (which neither incentivizes recruitment directly nor charges an entry fee) maximizes the SO's expected profit. That is, an SO who offers the possibility of focusing on retail sales is better off not charging an entry fee or directly compensating distributors for recruiting new distributors into their down-line.

5.2   Non-common prior beliefs about the number of entrants

We now examine the case in which different agents may hold different prior beliefs about the number of potential entrants. We relax the assumption that the agents' discount factors are identical as follows. For every time $t \in \mathbb{N}$, conditional on the game reaching time t, each agent $i \in I$ holds a belief that the game continues for another period with a probability of δ_i, and that time t is the last period of the game with probability $1 - \delta_i$. Assume that the agents' discount factors are drawn from a distribution H with a support $[\underline{\delta}, \bar{\delta}]$, where $\bar{\delta} < 1$. The higher a player's discount factor is, the more optimistic he is about the number of future entrants.

Since our agents now have different prior beliefs, we must be careful when we choose the solution concept as SPE and backward induction need not coincide. In an SPE, players correctly predict subsequent players' actions. However, when players put themselves in their opponents' shoes they might err when they hold prior beliefs that are different from their opponents' beliefs (for a comprehensive discussion of this issue in the context of time preferences, see Sarafidis, 2006). For the sake of brevity, we shall use SPE in order to solve the model. However, the next result holds under both concepts.

Proposition 9 *Let $q = 0$. There exists no incentive-compatible reward scheme R such that in its induced game there exists an SPE in which the SO makes a strictly positive expected profit.*

Proof. Consider an agent j who enters the game in round t. Agent j's expected number of descendants (in G) is smaller than $\sum_{i=1}^{\infty} \frac{\bar{\delta}^i}{t+1}$. As this expression goes to 0 when t tends to infinity, j's expected number of descendants (in G) goes to 0 as well. Incentive compatibility implies that $R(T_i) \leq \phi_R(|G_i| - 1)$. Hence, there is a round $t^\star$ from which time onwards purchasing a license cannot be a best response regardless of an agent's beliefs about his successors' actions. In an SPE, agents correctly predict that their opponents will never purchase a license after round $t^\star$. Therefore, we can use a standard backward induction argument to show that, in an SPE, agents never purchase a license (independently of whether they use their opponents' discount factors or put themselves in their opponents' shoes and use their own discount factor). ∎

The difference between ABEE and non-common priors

Players 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 players' evaluation of the future is affected by the overall behavior of their opponents and, in particular, by the events that take place at the beginning of the game. It is this history-dependent evaluation that enables a pyramid scam to work.

6 Related Literature

Throughout most of the analysis, we used ABEE as a solution concept. This concept was developed in Jehiel (2005) and later extended in Jehiel and Koessler (2008). In an ABEE, agents bundle contingencies into analogy classes and form beliefs that are measurable with respect to these analogy classes. Other prominent models in which agents reason in terms of a coarse representation of the correct underlying distribution are Piccione and Rubinstein (2003), Guarino and Jehiel (2011), Eyster and Piccione (2013), and Steiner and Stewart (2015).

A 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 their opponents' actions depend on their opponents' private information. In the context of social learning, Eyster and Rabin (2010) study a model in which agents best respond to the belief that their opponents' actions are conditioned on their private information and do not reflect learning from these opponents' predecessors' behavior. This type of play can be viewed as level-1 thinking in a cognitive hierarchy model in which the level-0 type best responds to his private information without

taking his opponents' actions into account. The main conceptual difference from our level-k thinking model is that in our model, the level-0 type's beliefs are determined in equilibrium.

Our paper includes a contracting interaction between a fully rational monopolist and boundedly rational agents. Spiegler (2011) offers a textbook treatment of models in which "rational" firms interact with boundedly rational agents. In Eliaz and Spiegler (2006, 2008), a principal interacts with agents who differ in their ability to predict their future tastes. In the context of auctions, Crawford et al. (2009) illustrate how agents who are characterized by level-k thinking can be exploited by an auctioneer who uses an "exotic" auction. Jehiel (2011) studies feedback design in auctions when bidders bundle their opponents' behavior (possibly in different auction formats) into analogy classes.

Our work also relates to the literature on bubbles and speculative trade. Bianchi and Jehiel (2010) analyze a model of speculative bubbles in which there are two types of investors. While rational investors are aware of the subtleties of the equilibrium patterns, partially sophisticated investors expect constant buy/sell orders that match the actual frequencies averaged over time. Eliaz and Spiegler (2007, 2009) apply a mechanism-design approach to speculative trade between agents who hold different prior beliefs.

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 the bubble started* and how many other agents are aware of the bubble. Abreu and Brunnermeier show that the bubble may persist long after all of the investors are aware of its existence.

As in the case of a bubble, time plays a major role in pyramid scams as, potentially, agents who arrive early have an incentive to join the scam since they can benefit at the expense of those who join later. However, uncertainty about arrival time *cannot lead to participation* in a pyramid scam. The reason that such uncertainty is sufficient to maintain a bubble in Abreu and Brunnermeier's model is that, unlike a typical pyramid scam, their trading game is not a zero-sum game. Underlying the bubble in the model there is an exogenous process that represents behavioral traders. When the rational investors ride the bubble in their model, they make positive profits at the expense of those behavioral traders.

Pyramids and multilevel marketing schemes have received some attention outside of the economic literature. Emek et al. (2011) characterize mechanisms for multilevel

marketing and find that geometric schemes are the only schemes that satisfy child dependence (the root's reward is uniquely determined by the rewards of its children), depth-level dependence (the root's reward is uniquely determined by the number of nodes on each level of the down-line), and additivity (which pins down the linear structure). Gastwirth (1977) and Bhattacharya and Gastwirth (1984) use the random recursive tree model to demonstrate the fraud underlying two real-world pyramid schemes: only a small fraction of the participants can make profits that cover the entry fees. In these two papers there are no strategic elements.

Relation to the centipede game

There is some resemblance between the structure of pyramid scams and the famous centipede game. 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 all *continue* for a few rounds. Jehiel (2005) resolves this paradox by showing that it is possible to sustain an ABEE in which players *continue*. In fact, under the coarsest partition, in a pure-strategy ABEE, players either always *stop* or always *continue* (except in the last round).

Let us understand the difference between pyramid scams and the centipede game in the context of analogy-based expectations. In the centipede game, in every history that is reached with positive probability there is *at most one node in which an agent stops*. This implies that in histories that affect the agents' analogy-based expectations, the frequency of *continue* is much larger than the frequency of *stop*. Moreover, if any of the players continues in an ABEE, then at least one of them believes that his opponent always continues, and this leads to the extreme result described above. In a pyramid scam, agents continue to reject offers long after the first rejection, and this affects (negatively) the agents' analogy-based expectations that their opponents will join the scam. The negative effect makes it significantly harder for the SO to sustain a pyramid scam. In fact, due to this effect, in some instances (e.g., when p is large), even if the game is arbitrarily long, it is impossible to find a payoff structure that would induce an ABEE in which agents participate in a pyramid scam.

7 Concluding Remarks

We presented a model that can be used both to study legitimate multilevel marketing and to understand the conditions that allow pyramid scams to work. We showed that agents who use simplified representations of their opponents' behavior may participate

in a pyramid scam, and examined the characteristics of reward schemes that underlie such scams. Our results provided a rationale for why such schemes (as well as legitimate multilevel marketing schemes) often compensate participants based on multiple levels of their down-lines. We connected strategic sophistication and vulnerability to pyramid scams, and showed that in some instances where a population of agents who use simple theories of their opponents' behavior is not vulnerable to pyramid scams, a population of agents whose degree of sophistication is higher will participate in such scams. At the methodological level, we presented a level-k thinking model in which the naive behavior is determined in equilibrium. Finally, we contributed to the industrial organization literature by characterizing the optimal multilevel-marketing scheme when agents are fully rational.

The agents' arrival process

The process we used throughout the paper is borrowed from the applied statistics literature, where it is referred to as the *uniform random recursive tree model*. We allowed the number of rounds to be random, whereas most of the uniform random recursive tree model's applications assume that there is a fixed number of n rounds. All of our results hold under the latter specification; however, we chose not to apply it in order to make the model more realistic. In fact, relaxing the randomness assumption would allow us to prove additional, stronger results. In Appendix B, we assume that it is commonly known that the game lasts for n rounds. This simplification allows us to provide a flavor of of the size of an optimal pyramid scam when $q = 0$.

Legal restrictions

In some countries, using recruitment-based compensation and charging entry fees are criminal offenses (see, e.g., Competition Bureau, 2015). Let us incorporate these restrictions into our model and study the possibility of sustaining a pyramid scam when $\phi_R = 0$ and $a_R(\tau) = 0$ for every $\tau \in \mathbb{N}$. Instead of an entry fee, assume that the SO is able to force each agent who wants to become a distributor to purchase initial stock of x units of the good at a price of η_R per unit. Consider the $q = 0$ case. There is a one-to-one mapping between this specification and the model we presented in the previous sections with $\phi_R = x\eta_R$ and $a_R(\tau) = xb_R(\tau)$ for every $\tau \in \mathbb{N}$. Thus, preventing the SO from charging explicit fees and from paying recruitment-based compensation would not affect his ability to sustain a pyramid scam.

References

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

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

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

[4] Bort, R. (2016): "John Oliver Says Multilevel Marketing Companies Like Herbalife are Pyramid Schemes," *Newsweek,* Nov. 7. http://www.newsweek.com/john-oliver-last-week-tonight-herbalife-pyramid-scheme-517881.

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

[6] Competition Bureau (2015): "Multi-level Marketing and Pyramid Selling," Nov. 15. Retrieved from http://www.competitionbureau.gc.ca.

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

[8] Direct Selling Association (2016): "Growth and Outlook Report: U.S. Direct Selling in 2015." Retrieved from http://www.dsa.org/benefits/research/factsheets.

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

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

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

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

[13] 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.

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

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

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

[17] Federal Trade Commission (2012): "FTC Action Leads to Court Order Shutting Down Pyramid Scam: Thousands of Consumers Burned by BurnLounge." Retrieved from https://www.ftc.gov/news-events/press-releases/2012/03/ftc-action-leads-court-order-shutting-down-pyramid-scamthousands.

[18] Federal Trade Commission (2014): "FTC Settlement Bans Pyramid Scheme Operators from Multi-Level Marketing." Retrieved from https://www.ftc.gov/news-events/press-releases/2014/05/ftc-settlement-bans-pyramid-scheme-operators-multi-level.

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

[20] Federal Trade Commission (2016b): "Consumer Information: Multilevel Marketing." Retrieved from https://www.consumer.ftc.gov /articles/0065-multilevel-marketing.

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

[22] Guarino, A. and Jehiel, P. (2011): "Social Learning with Coarse Inference," *American Economic Journal: Microeconomics,* 5, 147–174.

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

[24] Hoyle, A. (2016): "Can you Really Earn 350,000 a Year Selling Aloe Vera? As Thousands of Women Join a Controversial Home-Selling Empire, Read Candice's Very Cautionary Tale," *Daily Mail,* May. 19. http://www.dailymail.co.uk/femail/article-3598017/Can-really-earn-350-000-year-selling-aloe-vera-thousands-women-join-controversial-home-selling-empire-read-Candice-s-cautionary-tale.html.

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

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

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

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

[29] McCarthy, E. (2016): "Dear Facebook Friends, Please Stop Trying to Sell me Stuff. Its Getting Awkward.," *The Washington Post,* Feb. 185. https://www.washingtonpost.com/news/the-intersect/wp/2016/02/18/why-your-facebook-feed-is-drowning-in-dreadful-sales-pitches/.

[30] McCrum, D. (2014): "The Multi-Level Endgame, Part One: The FTC," *Financial Times,* Feb. 10. https://ftalphaville.ft.com/2014/02/10/1753272/the-multi-level-endgame-part-one-the-ftc/.

[31] McKown, C. (2017): "Legit Business Opportunity or Pyramid Scheme? Six Signs to Watch Out for if Youre Presented with a Multilevel Marketing Job," *CNBC,* Jun. 09. https://www.cnbc.com/2017/06/09/pyramid-scheme-or-legit-multilevel-marketing-job.html.

[32] Milgrom, P. and Stokey, N. (1982): "Information, Trade and Common Knowledge," *Journal of Economic Theory,* 26, 17–27.

[33] Moris, S. (1996): "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics,* 110, 1111–1133.

[34] "Multi-level Marketing in America: Pharaonic Creations. The Growing Battle over how to Spot a Pyramid Scheme", *The Economist,* Oct. 24. 2015. https://www.economist.com.

[35] Nocera, J. (2015): "The Pyramid Scheme Problem," *The New York Times*, Sep. 15. https://www.nytimes.com/2015/09/15/opinion/joe-nocera-the-pyramid-scheme-problem.html?.

[36] 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.

[37] Parloff, R. (2016): "Herbalife Deal Poses Challenges For The Industry," *Fortune*, Jul. 19. http://fortune.com/2016/07/19/herbalife-deal-challenges-industry/.

[38] Sarafidis, Y. (2006): "Games with Time Inconsistent Players," *SSRN Working Paper.*

[39] Schuette, B. (n.d.): "Multi-Level Marketing or Illegal Pyramid Scheme?" Retrieved from http://www.michigan.gov/ag/.

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

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

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

[43] Wieczner, J. (2017): "Herbalife Paid a 200 Million Fine. Then the FTC Screwed It Up," *Fortune*, Feb. 2. https://www.fortune.com.

8   Appendix A

A proof that the derivative of (13) with respect to p is negative

Observe that the derivative of (12) is

$$v'(p) := v' = l_2 + 2pl_3 + 3p^2 l_4 + \dots \tag{14}$$

Let us derive the x-th component of (13), given by $\frac{l_x(1+pv)^{x-1}}{(1+v)^x}$, with respect to p to

obtain

$$\frac{l_x}{(1+v)^{2x}}[(1+v)^x(x-1)(1+pv)^{x-2}(v+v'p)] \tag{15}$$

$$-\frac{l_x}{(1+v)^{2x}}[x(1+v)^{x-1}(1+pv)^{x-1}v']$$

Observe that

$$v+v'p = l_1 + 2pl_2 + 3p^2l_3 + 4p^3l_4 + ... \tag{16}$$

and that for $x=1$, $\frac{l_x}{(1+v)^{2x}}[(1+v)^x(x-1)(1+pv)^{x-2}(v+v'p)] = 0$. Hence, we can reduce the lower part of (15) for the x-th component of (13) from the upper part of (15) for the $(x+1)$-th component of (13) and sum over all $x \geq 1$ to obtain the derivative of (13) with respect to p:

$$\sum_{x=1}^{\infty}\frac{x(1+pv)^{x-1}}{(1+v)^{x+1}}[l_{x+1}(v+v'p) - l_x v']$$

This derivative equals the sum of the elements of the following infinite square matrix:

$$\begin{pmatrix}
(l_2l_1 - l_1l_2)\gamma_{11} & (l_2l_2 - l_1l_3)\gamma_{12} & (l_2l_3 - l_1l_4)\gamma_{13} & (l_2l_4 - l_1l_5)\gamma_{14} & \cdots & \vdots \\
(l_3l_1 - l_2l_2)\gamma_{21} & (l_3l_2 - l_2l_3)\gamma_{22} & (l_3l_3 - l_2l_4)\gamma_{23} & (l_3l_4 - l_2l_5)\gamma_{24} & \cdots & \vdots \\
(l_4l_1 - l_3l_2)\gamma_{31} & (l_4l_2 - l_3l_3)\gamma_{32} & 0*\gamma_{33} & (l_4l_4 - l_3l_5)\gamma_{34} & \cdots & \vdots \\
(l_5l_1 - l_4l_2)\gamma_{41} & (l_5l_2 - l_4l_3)\gamma_{42} & (l_5l_3 - l_4l_4)\gamma_{43} & 0*\gamma_{34} & \cdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots
\end{pmatrix}$$

where

$$\gamma_{xy} = \frac{xyp^{y-1}(1-pv)^{x-1}}{(1+v)^{x+1}}$$

Observe that $\gamma_{xy} > \gamma_{yx}$ if and only if $x > y$. Hence, the sum of the matrix's cells is strictly negative if

$$\frac{l_2}{l_1} > \frac{l_3}{l_2} > \frac{l_4}{l_3} > ... \tag{17}$$

We now prove (17). With a slight abuse of notation, let k denote an agent who enters

38

the game in round k. Denote by $l_{j,t}$ the expected number of agents in the j-th level of G_k in the case where the game ends at the end of round $t > k$. Suppose that $l_{1,t} > 0$ such that $\frac{l_{2,t}}{l_{1,t}}$ is well defined. An increase in t raises $\frac{l_{2,t}}{l_{1,t}}$ since $l_{1,t} - l_{1,t-1} = \frac{1}{t}$ and $l_{2,t} - l_{2,t-1} = \frac{1}{t} \left(\frac{1}{k+1} + ... + \frac{1}{t-1} \right)$. Since

$$\frac{l_{j+1,t}}{l_{j,t}} = \frac{l_{j+1,t-1} + \frac{1}{t} l_{j,t-1}}{l_{j,t-1} + \frac{1}{t} l_{j-1,t-1}} \tag{18}$$

it follows that the LHS of (18) is increasing in t whenever it is well defined. Observe that (18) also implies that

$$\frac{l_{j+1,t}}{l_{j,t}} \leq max \left\{ \frac{l_{j+1,t-1}}{l_{j,t-1}}, \frac{l_{j,t-1}}{l_{j-1,t-1}} \right\} \tag{19}$$

By the previous arguments,

$$\frac{l_{j+1,t-1}}{l_{j,t-1}} < \frac{l_{j+1,t}}{l_{j,t}} \leq \frac{l_{j,t-1}}{l_{j-1,t-1}} < \frac{l_{j,t}}{l_{j-1,t}} \tag{20}$$

which concludes the proof.

9 Appendix B: Fixed Number of Rounds

Let us simplify our model by relaxing the uncertainty regarding the number of future entrants and focusing on the $p = 0$ case, which is the most favorable one for sustaining a pyramid scam according to Proposition 6. Formally, instead of having a fixed discount factor δ, we assume that it is commonly known that the game lasts for n rounds. This simplification allows us to illustrate the profit-maximizing scheme. It should be mentioned that our previous results hold under minor modifications. For instance, in Proposition 6, we need to change the statement "for every $\delta > \delta^\star$" to "for every $n \geq 32$."

We now solve for the incentive-compatible reward scheme R that maximizes the SO's expected profit (as before, we consider the ABEE that guarantees the SO the highest profit among all of the ABEEs in the scheme's induced game) when $q = 0$ and $n = 100$.

Our agents are risk neutral and, in an ABEE, they hold beliefs that are different from those of the SO. This resembles the case of a bet between risk-neutral agents who hold different prior beliefs. In such a case, it is well known that these agents always prefer to increase the size of the bet. Therefore, to discuss profit maximization we must

bound the size of the bet. We shall assume that each agent has a budget $B > 0$ that he cannot exceed (i.e., $\phi_R \leq B$) and that this budget is large compared to c (otherwise it would be impossible to make these agents take part in a pyramid). To economize on notation, we set $B = 1$ and fix $c = 0$. The risk neutrality of the agents along with their budget pin down the entry fee as it must be equal to 1 at the optimum (otherwise it would be possible to raise the SO's profit by scaling up all of the commissions and the entry fees).

The profit-maximizing incentive-compatible reward scheme R is a collection of commissions per level $a_R(1), ..., a_R(n)$ such that $1 \geq a_R(\tau)$ for every level τ. In the ABEE of $\Gamma(R)$ in which the SO's profit is maximized, there is a last round $k < n$ in which an agent is willing to purchase a license. Denote this round by k_R and denote by L_t^h the expected number of distributors in the h-th level of the t-th entrant's downline according to his analogy-based expectations. Since R maximizes the SO's profit, $\sum_{h=1}^{n} L_{k_R}^h a_R(h) = 1$. That is, the last agent to purchase a license is indifferent between purchasing a license and not doing so (otherwise it would be possible to scale down the commissions and increase the SO's expected profit).

Let us examine the SO's strategy. If, given a scheme and a profile of strategies, it is a best response for the SO to make an offer in round $t < k_R$, it is also a best response for the SO to make an offer in round $t + 1$ since the expected commissions paid to a distributor weakly decrease with his time of entry. Hence, when we compare different schemes, we can focus on ABEEs in which the SO makes offers in a subset of *consecutive* rounds.

The next step is to show that, without loss of generality, we can restrict attention to profiles in which the SO makes offers only in the *first k_R* rounds. Consider an ABEE in which the SO makes offers in rounds $t, ..., k_R$, where $t > 1$. The number of distributors in T is a random variable that can take values in $\{0, ..., k_R - t + 1\}$. Let $t^\star$ be a value that maximizes the SO's expected profit given the scheme R. Consider a profile of strategies σ in which agents purchase licenses in rounds $1, ..., t^\star$ and the SO makes offers in these rounds (in σ, agents reject offers made in rounds $t' > t^\star$). The SO's expected net profit increases due to this transition, but the new profile may not be an ABEE. It is possible to calculate L_k^h and $L_{t^\star}^h$ for every $h \in \{1, ..., n\}$ and show that $L_k^h \leq L_{t^\star}^h$. This implies that the SO can fix the entry fees, scale down the commissions at each level, and make the new profile an ABEE. Hence, in the context of profit maximization, we can restrict our attention to ABEEs in which the SO makes offers in the first k_R rounds.

It is left to set the different commissions such that the SO's actual payments minus

the fees he can charge are minimized. We need to solve this problem for every potential $k_R \leq n$ and maximize over these k_R's. Denote the j-th entrant to T by i_j. The solution requires us to compare, for every $\tau \in \{1, ..., n\}$, the payments that the SO can extract from the agents based on $a(\tau)$, that is, $k_R L_{k_R}^\tau a(\tau)$, with the payments that he actually has to pay based on $a(\tau)$, that is, $\sum_{j=1}^{k_R} Pr[d_G(SO, i_j) \geq \tau + 1]a(\tau)$. The latter expression can be calculated due to a result in Mahmoud (1991): for $i_j \geq x + 1$,

$$Pr[d_G(SO, i_j) = x] = \frac{1}{(i_j - 1)!} \begin{bmatrix} i_j - 1 \\ x \end{bmatrix} \tag{21}$$

where $\begin{bmatrix} i_j - 1 \\ x \end{bmatrix}$ is the signless Stirling number of the first kind (the number of permutations of i_j elements with x disjoint cycles). We can write $L_{k_R}^\tau$ explicitly as

$$\frac{\sum_{i_1 = k_R + 1}^{n - (\tau - 1)} \sum_{i_2 = i_1 + 1}^{n - (\tau - 2)} \cdots \sum_{i_\tau = i_{\tau - 1} + 1}^{n} \frac{1}{i_1 i_2 \ldots i_\tau}}{\left(1 + \sum_{j = k_R + 1}^{n} \frac{1}{j}\right)^\tau} \tag{22}$$

We can now calculate the profit-maximizing scheme given a fixed k_R (we must make this calculation for every $k_R \leq n$). The incentive-compatible scheme that maximizes the SO's profit must assign weights to the different a's such that:

- For every $\tau \in \{1, ..., n\}$, $a_R(\tau) \leq 1$.

- $\sum_{\tau=1}^{n} L_{k_R}^\tau a_R(\tau) = 1$.

- The size of the different $a_R(\tau)$'s is decided according to the following cost-benefit ratio:

$$\frac{\sum_{i=1}^{k_R} Pr[d_G(SO, i) \geq \tau + 1]}{k_R L_{k_R}^\tau} \tag{23}$$

We start with the $a(\tau)$'s for which the ratio is the lowest and increase them until one of the following occurs: either $a(\tau) = 1$ or $\sum_{\tau=1}^{n} L_{k_R}^\tau a_R(\tau) = 1$. We continue this procedure until $\sum_{\tau=1}^{n} L_{k_R}^\tau a_R(\tau) = 1$.

The scheme and the ABEE that maximize the SO's profit are such that 3 agents purchase a license ($k_R = 3$), $a_R(1) = 0.814, a_R(2) = ... = 1$, and the SO's profit is 1.884. With only two layers of compensation, the scheme and ABEE that maximize the SO's expected profit have $k_R = 3$, $a_R(1) = 0.922$, $a_R(2) = 1$, and the SO's maximal

41

profit is 1.758.