

Customer Referral Incentives and Social Media*

Ilan Lobel, Evan Sadler[†] and Lav R. Varshney[‡]

March 23, 2015

Abstract

We study how to optimally attract new customers using a referral program. Whenever a consumer makes a purchase, the firm gives her a link to share with friends, and every purchase coming through that link generates a referral payment. The firm chooses the referral payment function and consumers play an equilibrium in response. The optimal payment function is nonlinear and not necessarily monotonic in the number of successful referrals. If we approximate the optimal policy using a linear payment function, the approximation loss scales with the square root of the average consumer degree. Using a threshold payment, the approximation loss scales proportionally to the average consumer degree. Combining the two, using a linear payment function with a threshold bonus, we can achieve a constant bound on the approximation loss.

1 Introduction

Referrals have emerged as a primary way through which companies in the social networking era acquire new customers. Instead of spending money on traditional advertising,

*We thank IBM (OCR grant) for its support of this research project.

[†]Stern School of Business, New York University – {ilobel, esadler}@stern.nyu.edu

[‡]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign – varshney@illinois.edu

many companies now rely on social media-based referral programs to bring in new customers. Referral programs are used throughout the economy, but their use is particularly common among start-ups, where they are seen as both an affordable and an effective way to grow.

Referral programs come in many forms. Some companies pay customers for every new referral they bring in. Others pay customers only if they bring in sufficiently many referrals. We propose a framework to analyze different referral program designs and determine good payment rules. There are two main parts to our analysis. First, we identify the optimal payment function; for each possible number of successful referrals that a consumer makes, we find the optimal reward the firm should pay. Second, we examine simple and easy-to-implement payment functions—linear, threshold, and a combination of the two—to assess how well they can approximate the profits of the optimal policy.

Two companies that exemplify the success of referrals as a customer acquisition strategy are Living Social and Dropbox. Living Social is a daily deal website that offers discounted prices for a variety of goods and services. Living Social was launched in 2007 and achieved a multi-billion dollar valuation only four years later. At the heart of Living Social’s rapid growth was its viral marketing strategy. Whenever someone buys a deal from them, Living Social encourages the customer to post the transaction on social media. If at least three friends subsequently purchase based on this post, Living Social provides the customer with a full refund of the purchase price.

Dropbox, a cloud-based file hosting service that was founded in 2008, also achieved extraordinarily fast growth through referrals. Dropbox offers every user 2GB of free storage, and it increases this amount by 500MB for every friend the user refers (or 1GB if a friend signs up for a professional account), up to a maximum of 16GB for free. The CEO of Dropbox attributes much of the company’s early success, going from a hundred thousand to four million users in just 15 months, to its referral program [Houston, 2010]. Many other technology companies have grown through referral programs, including the electronic payments company PayPal in the early aughts, the Living Social competitor Groupon, and the car service companies Uber and Lyft.

The key question guiding our analysis is: what is the best way to structure referral payments? Linear payments, whereby every successful referral generates the same payment to the customer making referrals, are the simplest and probably the most common model. Dropbox uses linear payments capped at 28 referrals. Uber uses a linear payment function, with a payout of twenty dollars per successfully referred friend. In contrast, Living Social uses a threshold model in which two referrals yield no benefit, but three lead to a full refund. In principle, the optimal referral payments could follow a linear function, a threshold function, or any other nonlinear payment function.

We consider a firm that designs its referral incentive program with two objectives in mind: extract immediate revenue and advertise to potential customers. We focus on the interaction between this firm and a focal consumer who is currently deciding whether to purchase the firm's product. If the consumer buys the product, she has the option to refer as many of her friends as she wants. The firm wishes to maximize a linear combination of the revenue obtained from the focal consumer and the value obtained from her referrals. The consumer plays an equilibrium of a network game in which she balances the costs and rewards associated with making referrals. The consumer takes into account the value of the referral program to her friends when she calculates the probability that her friends will purchase the firm's product. When a referred friend purchases the firm's product, we say that a conversion occurs.

An important assumption is that the firm values referrals, both those that convert and those that do not, but only pays consumers for conversions. We make this assumption because the advertising value of a referral program goes beyond immediate sales, including product and brand awareness, as well as future sales. Even though the firm values referrals, we assume that the firm chooses to pay consumers only for the conversions they generate. One reason is that measuring referrals is more difficult than measuring purchases. More significantly, paying for referrals directly creates opportunities to cheat. Customers could send referrals to fake e-mail accounts, recouping the cost of their purchases without creating any value for the firm. Requiring a purchase for the referral to count towards a reward

avoids this problem. We consider the case in which conversions and non-conversions are valued differently in Section 6.1.

Our first main result characterizes the optimal referral policy when there are no restrictions on the payment function the firm can use. A key idea in our analysis is a decomposition of the expected referral payment into the portion that compensates consumers for the cost of making referrals and the remainder, which we interpret as the real reward or as a discount that the consumer receives on the purchase price. Since they can bring in more referrals, high-degree consumers will optimally receive larger rewards, making them more likely to purchase. Although the optimal expected rewards are increasing in the consumer’s degree, the optimal payment function itself can be non-monotonic in the number of conversions. This non-monotonicity occurs because the firm values referrals but pays for conversions. The number of conversions is a stochastic function of the number of referrals, so the expected reward a consumer receives is a smoothed version of the payment function, and precise control of the expected reward may require a complex payment rule. This implies that the optimal policy is, at least in some instances, quite complicated and potentially impractical to implement. Few if any firms would ever offer a deal like “bring 3 friends and earn \$20, or bring 4 friends and earn \$15” as consumers would likely find such programs confusing and would have difficulty responding optimally.

The non-monotonicity of optimal payments motivates our study of simpler referral functions that companies could realistically offer. We measure the performance of these payment functions by benchmarking them against the returns from using the optimal nonlinear payments. An obvious starting point is to consider linear payment functions. To evaluate how good of an approximation we can achieve with a linear policy, we perform a worst-case analysis with respect to the degree distribution of the network. We show that worst-case losses from using the best linear incentive functions scale according to the square root of the average degree of the social network. Another natural approximation is a threshold payment function, though these generally fare much worse than linear policies. With a threshold payment function, the worst-case losses relative to the optimum scale linearly with the average

degree in the network.

The different scaling rates for linear and threshold policies give an incomplete comparison because losses in each case come from very different sources. When the degree distribution has a heavy tail, linear policies are prone to overpaying high-degree consumers. Threshold policies provide a cap on payments, but they have difficulty properly compensating consumers for the costs of making referrals. To render these insights more concretely, we present a numerical analysis of a few special cases. Comparing the performance of a linear payment function across Poisson, geometric, and power-law degree distributions, we can clearly see how tail thickness negatively impacts profits. In social media settings, in which degree distributions typically follow a power law and the marginal cost of making a referral is low, threshold payments perform relatively well. In fact, if the marginal cost of making a referral is zero, we can show that the losses from using a threshold payment are bounded by a constant across all possible degree distributions.

Since these two types of policies have complementary strengths, combining them may offer a significant improvement over either one alone. Making linear payments, with an added threshold bonus, captures the best features of both policies. The linear part compensates consumers for the social costs of making referrals, and the threshold bonus can provide the desired discount for high-degree consumers without overpaying. This small amount of added complexity allows us to obtain a constant bound on the approximation loss, regardless of the degree distribution and regardless of consumer costs for making referrals.

1.1 Related Literature

Personal referrals have a powerful effect on consumers. Marketing researchers, using the term “word-of-mouth,” have studied the impact of referrals for decades. This work credits interpersonal communication with far greater influence over consumer attitudes and behavior than either conventional advertising or neutral print sources [Buttle, 1998], and the value of personal referrals can constitute a significant portion of a customer’s value [Kumar et al., 2007]. With the growth of online platforms, and social media in particular,

we can now measure these effects more precisely than ever before. For instance, Godes and Mayzlin [2004] are able to measure word-of-mouth effects using conversations from an online discussion forum. Studies in both the computer science literature and the information systems literature draw on data from online social networks to assess just how much people influence one another in different settings [Leskovec et al., 2007, Aral and Walker, 2011]. There is also work in experimental economics showing that social learning effects on demand are at least as big as traditional advertising channels (see [Mobius et al., 2011]).

Parallel work asks questions about design. Algorithmic and simulation-based approaches allow researchers to study the relative importance of weak and strong ties [Goldenberg et al., 2001], optimal seeding strategies [Kempe et al., 2003], and how to use viral product features to maximize the spread of adoption in a network [Aral et al., 2013]. A number of papers in the economics literature explore word-of-mouth communication as a signaling game in which consumers learn about the quality of a product through friends, and the firm tries to manipulate the information consumers receive, either through hired “trendsetters” [Chatterjee and Dutta, 2014] or launch strategies [Campbell, 2015]. Campbell [2013] studies pricing in the presence of word-of-mouth learning. We study a context in which the firm more directly manipulates word-of-mouth information transmission through its referral program.

Our equilibrium analysis draws on the recent economics literature on games with local network externalities [Ballester et al., 2006, Sundararajan, 2007, Galeotti et al., 2010]. A referral program provides a way to artificially create network externalities, inducing complementarities in friends’ purchase decisions; an equilibrium of our consumer game is mathematically equivalent to an equilibrium of a network game with strategic complements. Candogan et al. [2012], Bloch and Quérou [2013], and Cohen and Harsha [2015] build on these results to create a theory of optimal pricing for a monopolist selling goods to networks of consumers. Our paper brings a new perspective to this growing literature, taking a mechanism design approach that provides insight on how to structure the network externality in order to generate the desired information transmission.

A few other papers take up the problem of referral payment design. Biyalogorsky et al.

[2001] consider how the firm can jointly optimize pricing and referral incentives over the lifetime of a consumer. We introduce two important innovations on this approach. First, our consumers strategically anticipate the value that their friends will derive from the referral program. Second, we address the role of the social network structure. Libai et al. [2003] study the somewhat different setting of affiliate marketing, allowing a richer information set and individualized referral contracts. As typical customer referral programs cannot implement individualized referral contracts or keep track of leads separately from conversions, we restrict our attention to pay-per-conversion systems.

Our approach necessarily abstracts away from many details of referral programs that are nevertheless important. In particular, the medium through which customers communicate with each other has implications both for the cost of making referrals and the value the firm derives from referrals. Chu and Kim [2011] note that the influence people exert on each other through online interactions is different than that through offline interactions, and Burke and Kraut [2014] present evidence from Facebook suggesting that one-on-one communication is fundamentally different than multicast communication. Although we do not explicitly address what medium is best for a referral program, our model does provide an appropriate framework to think about this question. The type of interactions we rely upon to generate referrals will determine the cost, payoff, and network parameters that are primitives in our model, allowing a comparison between optimal policies using different modes of communication.

2 The Referral Game

A firm produces a product at zero marginal cost and sells it at a fixed price $p > 0$. For now, we take the price as exogenous, but we address the issue of optimal pricing in Section 6.2. There is a focal consumer deciding whether to purchase at the current moment, and the focal consumer is situated in a network of other potential consumers. The firm uses a referral incentive program to expand its customer base and increase profits. When a consumer makes a purchase, she is offered a code or a link to share with her friends. When a friend uses that

code or link to make a purchase, the original consumer may receive a reward from the firm.

Each consumer has a value $v \in [0, 1]$ for the product drawn from a continuous distribution F , which has density f . A consumer has $d \in \mathbb{N}$ friends to whom she can send referrals, and d is drawn according to the degree distribution G . We use g_d to denote the probability that a consumer has d friends. The pair (v, d) represents a consumer's type. We assume valuations and degrees are drawn independently across all consumers, and the distributions F and G are common knowledge. We can think of the social network as an infinite tree, so the decision problem of one consumer looks essentially identical to that of any other consumer.¹

The firm chooses how to structure a referral program for the focal consumer. We describe the referral program as a payment function $\mathbf{w} : \mathbb{N} \rightarrow \mathbb{R}$ specifying a monetary payoff for each possible number of successfully referred friends. We assume that $\mathbf{w}(0) = 0$, and we represent the space of possible referral incentive functions by \mathcal{W} .

Each consumer who becomes aware of the product through a referral chooses a pair $\mathbf{s} = (b, r)$, where $b \in \{0, 1\}$ indicates whether the consumer purchases the product ($b = 0$ represents no purchase), and $r \in \{0, 1, 2, \dots, d\}$ represents the number of friends who are referred. Consumers incur a social cost $\delta \geq 0$ for each referral. The strategy space for a consumer with d neighbors is $A_d = \{(0, 0)\} \cup \{(1, r) \mid r = 0, 1, \dots, d\}$. A pure strategy for a consumer is a function $\sigma(v, d, \mathbf{w}) : [0, 1] \times \mathbb{N} \times \mathcal{W} \rightarrow A_d$ specifying an action for each possible valuation, degree, and payment function. We use b_σ and r_σ respectively to denote the two components of σ , and we use Σ to denote the space of pure strategies.

The firm's payoff depends on the actions of the focal consumer.² Let the random variable $N_\sigma(r)$ denote the number of neighbors who respond to a referral when the focal consumer refers r neighbors and other players use strategy σ . If the focal consumer chooses the pair

¹For our purposes, it is sufficient that a consumer's subjective beliefs about the decision problem her neighbors face are congruent with this structure. It is not necessary that the true network structure actually be an infinite tree.

²Alternatively, we could view this payoff in terms of the average behavior of a consumer who becomes aware of the product.

(b, r) , the firm earns

$$\pi_{(b,r)} = b(p - \mathbb{E}[\mathbf{w}(N_\sigma(r))] + ry),$$

where $y > 0$ is the value the firm places on a referral. We assume that $y > \delta$. If the consumer buys the product, the firm earns the purchase price p , less the expected referral payment, plus an additional payoff that is proportional to the number of referrals. The value y captures more than the potential value of an immediate purchase by the referred neighbor; it includes also the value of later purchases, future word-of-mouth, and increased brand awareness and brand loyalty. We think of y as the value of advertising generated through referrals. For most of what follows, we treat this value as independent of whether or not a referral results in a conversion, but in Section 6.1 we discuss what happens if the firm differentiates the two in its objective.

We consider equilibria of the consumer game in pure symmetric strategies, using σ' to denote a unilateral deviation when all other consumers play σ . The strategy σ induces expectations over the distribution of $N_\sigma(r)$. Playing σ' in response to others playing σ yields an expected payoff

$$u_{\sigma',\sigma}(v, d, \mathbf{w}) = b_{\sigma'}(v - p - \delta r_{\sigma'} + \mathbb{E}_\sigma[\mathbf{w}(N_\sigma(r_{\sigma'}))]). \quad (1)$$

A purchasing consumer earns her valuation for the product v , less the purchase price and the cost of making referrals, plus the expected referral payment. A strategy profile³ $\sigma \in \Sigma$ is a pure strategy Bayesian equilibrium of the consumer game if

$$u_{\sigma,\sigma}(v, d, \mathbf{w}) \geq u_{\sigma',\sigma}(v, d, \mathbf{w}) \text{ for all } (v, d, \mathbf{w}) \in [0, 1] \times \mathbb{N} \times \mathcal{W} \text{ and all } \sigma' \in \Sigma. \quad (2)$$

The firm first chooses a payment function \mathbf{w} . The focal consumer then observes this choice and plays a symmetric equilibrium for the consumer game in response. Our goal is to understand optimal or near-optimal payment functions in terms of the firm's value of a referral y , the consumer's cost of a referral δ , and the distributions F and G of consumer

³Note the slight abuse of notation: we use the same representation Σ to represent symmetric pure strategy profiles that we use to represent a consumer's space of pure strategies.

valuations and degrees. Formally, we study the firm’s optimization problem:

$$\begin{aligned} \sup_{\mathbf{w} \in \mathcal{W}, \sigma \in \Sigma} \quad & \mathbb{E}_{v,d} [b_\sigma (p - \mathbf{w}(N_\sigma(r_\sigma)) + r_\sigma y)] \\ \text{s.t.} \quad & u_{\sigma,\sigma}(v, d, \mathbf{w}) \geq u_{\sigma',\sigma}(v, d, \mathbf{w}) \quad \text{for all } \sigma' \in \Sigma, v \in [0, 1], d \in \mathbb{N}. \end{aligned} \tag{3}$$

We can interpret this model as a mechanism design problem, with the firm maximizing an objective subject to individual rationality and incentive compatibility constraints.

2.1 Remarks on the Model

Our modeling choices reflect a difference in perspective between the consumers and the firm. For a consumer deciding whether to buy, the decision problem is fundamentally static: she is making a single decision in a given moment. The potential for future purchases or future interactions with the firm do not enter the consumer’s thought process. Our focal consumer assumes that any friend she refers will face the same basic decision problem, including the referral program, so she accounts for the value of the referral program to her neighbors when she evaluates the value of the program to her. Even though friends’ purchase decisions occur in the future, the consumer game is formally equivalent to a one-shot simultaneous move game because consumers are symmetric and each makes only one decision.

At first glance, the assumption that the network rooted on the focal consumer is an infinite tree appears restrictive, but at least two interpretations suggest the model is more general. One interpretation focuses on consumers’ cognitive limitations: they adopt the infinite tree model as their subjective representation of the network because it simplifies decision-making. A fully Bayesian decision-maker situated in a network will have to solve a very complex decision problem before deciding whether to purchase and how many friends to refer. A second interpretation is that we have simply pruned the cycles from the network by eliminating extra links. In reality, any particular consumer may receive multiple referrals from different friends, but the firm would only attribute the referral to one of them. We can view our network as a subgraph in which we ignore links that produce redundant

referrals. This interpretation introduces the additional issue that some neighbors may not respond to referrals because they have already received one from another friend, or they have already learned about the product through other means. Accounting for this in the model, for instance making each neighbor already aware of the product with some independent probability, is straightforward and yields the same qualitative results.

In contrast to the consumer, the firm takes a broader view of the problem: the firm is interested in growing its business. Beyond the immediate purchase and referrals from the focal consumer, the firm is interested in future referrals to friends of friends, repeat purchases, the possibility of changing the referral program over time, and building brand awareness and loyalty. We abstract away from this complexity by incorporating all of these elements into the referral value y that we associate with the focal consumer. In principle, this value could emerge as a continuation value in a dynamic game in which the firm updates its referral program as the customer base grows, but adding this layer would not change our structural results for optimal and approximately optimal payment functions.

An alternative way to view the referral program is as a substitute for traditional advertising. We can think of the value y as an outside option representing the cost of achieving equivalent reach through standard advertising channels. Regardless of the interpretation, our representation of the firm’s problem provides a simple way to capture the important trade-offs in referral program design. Furthermore, our model can offer meaningful guidance even when information about the structure of the network is limited to the degree distribution.

3 Consumer Behavior

We first study the consumer equilibria for a fixed referral payment function $\mathbf{w} \in \mathcal{W}$. Throughout this section, we take \mathbf{w} as given, and we suppress dependence on \mathbf{w} in our expressions to simplify notation. Suppose the focal consumer believes that each neighbor will purchase in response to a referral independently with some probability P . This neighbor purchase probability, together with the number of neighbors the consumer has, determines the referral incentive’s value for the consumer. Let $B_P(r)$ denote a binomial random variable

with r trials and success probability P . If a consumer has d neighbors, each of whom purchases in response to a referral with probability P , the value of the referral program is

$$I_P(d) = \max_{r \in \{0, 1, \dots, d\}} \mathbb{E}[\mathbf{w}(B_P(r)) - r\delta]. \quad (4)$$

Combining equations (1) and (4), we see that a focal consumer with d neighbors will purchase if and only if her valuation v is at least $p - I_P(d)$. The threshold

$$v_P(d) = p - I_P(d) \quad (5)$$

is the lowest valuation at which a consumer with degree d will buy, given that a neighbor will purchase with probability P . If other players' strategies result in a neighbor purchase probability P , then the thresholds $v_P(d)$ define a best reply strategy. This best reply depends on the other players' strategies solely through the neighbor purchase probability P .

A random consumer, drawn according to the distributions F and G , who best responds to the neighbor purchase probability P will in turn purchase with probability

$$\phi(P) = \sum_{d \in \mathbb{N}} g_d (1 - F(v_P(d))). \quad (6)$$

The function ϕ allows us to give an alternative characterization of equilibrium strategy profiles. Instead of working directly with utility functions as in equation (2), we can work with neighbor purchase probabilities. There is a correspondence between equilibrium strategy profiles and fixed points of the map ϕ . Given an equilibrium strategy profile σ , neighbors purchase independently with some probability P_σ , and we must have $\phi(P_\sigma) = P_\sigma$. If P_σ is not a fixed point, then a random consumer best responding to P_σ would purchase with a different probability, which is inconsistent with P_σ being an equilibrium. Conversely, for any P^* that is a fixed point of ϕ , the thresholds $v_{P^*}(d)$ define an equilibrium strategy uniquely up to a measure zero set of valuations who are indifferent about purchasing. The correspondence between equilibria and fixed points of ϕ yields a simple existence proof.

Proposition 1. *A symmetric pure strategy Bayesian equilibrium of the consumer game exists.*

Proof. This amounts to showing that a fixed point exists for the function ϕ defined in equation (6). Since $I_P(d)$ is a continuous function of P , and F is a continuous distribution, the map ϕ is continuous. The existence of a fixed point is an immediate consequence of Brouwer's fixed point theorem. □

This result establishes existence but not uniqueness of the equilibrium. Some parameter values and payment functions support multiple neighbor purchase probabilities in equilibrium. For instance, suppose that $p > 1$, that $g_0 = 0$, and that $\mathbf{w}(d) > p$ for each $d > 0$. In this case, both $P_\sigma = 0$ and $P_\sigma = 1$ are equilibrium purchase probabilities. The former is an equilibrium in which no consumer purchases because the good is too expensive and the referral payment contributes no value since others never respond to referrals. The latter resembles a Ponzi scheme in which all consumers buy the good only because they expect referral payments to entirely cover the cost.

However, as long as the referral payments are not too generous, Ponzi scheme equilibria will not arise, and we should expect a unique consumer equilibrium. A sufficient condition for uniqueness is that $\phi(P) - P$ is a decreasing function of P . Equivalently, the equilibrium is unique if we have $|\phi(P_1) - \phi(P_2)| \leq |P_1 - P_2|$ for all $P_1, P_2 \in [0, 1]$. The following result provides one sufficient condition for uniqueness.

Proposition 2. *Let $\bar{w} = \sup_d (\mathbf{w}(d) - \mathbf{w}(d - 1))$, and let $\bar{f} = \sup_{v \in [0, 1]} f(v)$. If we have $\bar{f}\bar{w}\mathbb{E}_{D \sim G}[D] \leq 1$, then the pure strategy Bayesian equilibrium of the consumer game is unique.*

Proof. We compute

$$\begin{aligned} |\phi(P_1) - \phi(P_2)| &\leq \sum_{d \in \mathbb{N}} g_d \left| F(v_{P_2}(d)) - F(v_{P_1}(d)) \right| \leq \bar{f} \sum_{d \in \mathbb{N}} g_d |v_{P_2}(d) - v_{P_1}(d)| \\ &= \bar{f} \sum_{d \in \mathbb{N}} g_d |I_{P_1}(d) - I_{P_2}(d)| \leq \bar{f} \sum_{d \in \mathbb{N}} g_d \bar{w} |P_1 - P_2| \\ &= \bar{f}\bar{w}\mathbb{E}_{D \sim G}[D] |P_1 - P_2|. \end{aligned}$$

The last inequality is the critical step. A consumer with d neighbors can gain no more than

$\bar{w}|P_1 - P_2|$ per neighbor when the neighbor purchase probability changes. The conclusion follows. \square

The condition in Proposition 2 gives us a bound \bar{w} on the incremental benefit from making one more successful referral. To get an intuitive sense of how broadly the condition applies, suppose F is uniform so that $\bar{f} = 1$. In this case, the condition is satisfied if a consumer with an average number of neighbors can never obtain a total referral payment of 1, the highest amount that any consumer would be willing to pay for the product.

4 Optimal Incentives

We now turn to the firm's problem of selecting the optimal payment function \mathbf{w} , and we reintroduce the dependence on \mathbf{w} in our expressions. We can decompose the firm's referral payments into two components: part of the payment compensates consumers for making referrals, and the rest provides a reward (or an indirect discount) on the purchase price. If a consumer with d neighbors makes $r_\sigma(d)$ referrals in equilibrium σ , we can rewrite Eq. (4) as

$$\mathbb{E}_\sigma [\mathbf{w} (B_{P_{\sigma,\mathbf{w}}} (r_\sigma(d)))] = I_{P_{\sigma,\mathbf{w}}}(d) + r_{\sigma,\mathbf{w}}(d)\delta.$$

The value of the incentive $I_{P_{\sigma,\mathbf{w}}}(d) \geq 0$ to a consumer is equivalent to a reduction in the effective price, and $r_{\sigma,\mathbf{w}}(d)\delta$ is the amount the firm must pay in expectation to compensate the consumer for making referrals.

There are at least two reasons to use a referral incentive to discount the product beyond compensating for referrals. First, the referral incentive offers a way to price discriminate between consumers of different degrees. Since the firm's value for referrals y is larger than a consumer's cost of making a referral δ , consumers with more neighbors can generate more surplus if they make purchases. Intuitively, an optimal incentive should offer these consumers a lower effective price so that more of them buy. Second, if the good is subject to a Veblen effect, a referral incentive could provide a way to retain the demand associated with a high list price while discounting the good to an optimal effective price.

Decomposing referral payments in this way allows for a simpler representation of the firm's decision problem. The firm can fully describe consumer behavior using a sequence of referral and discount pairs $\{(r_d, I_d)\}_{d \in \mathbb{N}}$. For a consumer with d neighbors, the value r_d specifies how many referrals she makes conditional on a purchase, and I_d is the discount in equilibrium she obtains, which determines how likely the consumer is to make a purchase. The firm's problem is then to choose a sequence $\{(r_d, I_d)\}$ that will maximize profits. The firm is constrained to choose a sequence it can implement in some equilibrium of the consumer game. We can restate the firm's problem (3) as

$$\begin{aligned} \max_{\{(r_d, I_d)\}} \quad & \sum_{d \in \mathbb{N}} g_d (1 - F(p - I_d)) (p - r_d \delta - I_d + r_d y) \\ \text{s.t.} \quad & \{(r_d, I_d)\} \text{ implementable.} \end{aligned} \tag{7}$$

Implementability necessitates that the sequence of referral and discount pairs satisfies individual rationality and incentive compatibility constraints. Feasibility requires that $r_d \in \{0, 1, \dots, d\}$ and individual rationality mandates that $I_d \geq 0$ for each d . Since a consumer with many neighbors can always mimic one with fewer neighbors, incentive compatibility implies that:

- The r_d are non-decreasing in d , and whenever $r_d \neq r_{d-1}$ we have $r_d = d$;
- The I_d are non-decreasing in d , and $I_d > 0$ only if $r_d > 0$;
- Whenever $r_{d'} = r_d$, we have $I_{d'} = I_d$.

It turns out that satisfying the individual rationality and incentive compatibility constraints is sufficient to ensure that a sequence is implementable: for any sequence satisfying the above properties, there is a corresponding payment function \mathbf{w} and an equilibrium σ of the consumer game producing the given sequence.

Theorem 1. *A sequence of referral and discount pairs $\{(r_d, I_d)\}_{d \in \mathbb{N}}$ is implementable if and only if the above individual rationality and incentive compatibility constraints are satisfied.*

Proof. Given a sequence of discount pairs, an equilibrium neighbor purchase probability is uniquely determined as

$$P_{\sigma, \mathbf{w}} = \sum_{d \in \mathbb{N}} g_d (1 - F(p - I_d)).$$

Assuming this neighbor purchase probability, we can inductively construct a payment function \mathbf{w} that implements the corresponding referral and discount pairs. Note that the discount to customers of degree d can only depend on the values of $\mathbf{w}(k)$ for $k \leq d$.

For the case $d = 1$, if $r_1 = 0$ take $\mathbf{w}(1) = 0$, otherwise set $\mathbf{w}(1) = \frac{I_1 + \delta}{P_{\sigma, \mathbf{w}}}$. Suppose we have constructed a \mathbf{w} that implements a given sequence up to customers of degree d , and consider two cases. First, if $r_{d+1} \neq d + 1$, we have $r_{d+1} = r_d$ and $I_{d+1} = I_d$. Setting $\mathbf{w}(d + 1)$ sufficiently low will make the value of referring the $(d + 1)^{th}$ neighbor negative, so these consumers will mimic those of degree d as desired. Alternatively, we have $r_{d+1} = d + 1$ and $I_{d+1} \geq I_d$. We wish to define $\mathbf{w}(d + 1)$ so that

$$I_{d+1} + (d + 1)\delta = \sum_{j=0}^{d+1} \binom{d+1}{j} (P_{\sigma, \mathbf{w}})^j (1 - P_{\sigma, \mathbf{w}})^{d+1-j} \mathbf{w}(j).$$

Since we have already defined \mathbf{w} up through d , we can set

$$\mathbf{w}(d + 1) = \frac{1}{(P_{\sigma, \mathbf{w}})^{d+1}} \cdot \left[I_{d+1} + (d + 1)\delta - \sum_{j=0}^d \binom{d+1}{j} (P_{\sigma, \mathbf{w}})^j (1 - P_{\sigma, \mathbf{w}})^{d+1-j} \mathbf{w}(j) \right],$$

completing the proof. \square

Even with this relatively straightforward characterization of implementable sequences, the firm's problem as given by (7) is potentially difficult as the constraints are coupled through the consumer equilibrium. Rather than solving this problem directly, we consider first a relaxed problem in which we only impose the feasibility and individual rationality constraints. This allows us decouple each element of the sum, finding separately for each degree d the pair (r_d, I_d) that solves

$$\begin{aligned} \max_{r_d, I_d} \quad & (1 - F(p - I_d)) (p - r_d \delta - I_d + r_d y) \\ \text{s.t.} \quad & I_d \geq 0 \\ & r_d \in \{0, 1, \dots, d\}. \end{aligned} \tag{8}$$

The following theorem shows that the solution of the relaxed problem (8) satisfies our incentive compatibility constraints, and thus produces a solution for the original problem.

Theorem 2. *There is a solution to problem (8) in which $r_d = d$ for all d , and the I_d are non-decreasing in d . Consequently, this solution also optimizes problem (7). If we further have that ⁴*

$$(p - I) - \frac{1 - F(p - I)}{f(p - I)} \quad \text{is decreasing in } I, \quad (9)$$

then the optimal sequence $\{(r_d, I_d)\}_{d \in \mathbb{N}}$ is unique.

Proof. Since $y > \delta$, the firm should clearly pay for all available referrals, so we take $r_k = k$ for each k . The optimal I_k solves

$$\max_{I_d} (1 - F(p - I_d)) (p - I_d + d(y - \delta)),$$

subject to $0 \leq I_d \leq p$. The first derivative of the objective is

$$f(p - I_d) (p - I_d + d(y - \delta)) - (1 - F(p - I_d)).$$

Observe that for fixed I_d this derivative is increasing in d , so the value of I_d that attains the maximum is non-decreasing in d . If this derivative is positive (negative) for all $I_d \in [0, p]$, the unique solution is $I_d = p$ ($I_d = 0$). If Eq. (9) holds, and we have an interior optimum, the solution is the unique value

$$I_d = \sup \left\{ x : \frac{1 - F(p - x)}{f(p - x)} - (p - x) \leq d(y - \delta) \right\}. \quad (10)$$

□

The theorem above demonstrates that implementability does not meaningfully restrict the referral and discount pairs we can offer to consumers. We can achieve perfect price discrimination along consumer degrees and decouple the optimization for each part of the sum.

⁴Note that Eq. (9) is the classical regularity of the virtual values condition from Myerson [1981]’s seminal paper on auction design.

We can better understand the optimal policy if we consider the effective price $p - I_d$ consumers of degree d face. Take $d = 0$ in equation (10), and note that

$$p - I_0 - \frac{1 - F(p - I_0)}{f(p - I_0)} = 0$$

corresponds to the standard monopoly price solution, i.e., virtual value equals zero. For every additional neighbor a consumer has, the value of I_d increases, indicating a lower effective price.⁵ The more neighbors a consumer has, the greater the incentive to tradeoff immediate profits for additional referrals. Moreover, Eq. (10) tells us how to divide the added surplus $y - \delta$ between the consumer and the firm: the consumer pockets the direct change in I_d , while the firm earns the corresponding change in the ratio $\frac{1 - F(p - I_d)}{f(p - I_d)}$.

The practicality of optimal referral payments depends on the ease with which we can specify them and the ease with which consumers can respond to them. In general, the optimal payment function faces challenges on both counts because it will be non-monotonic in the number of successful referrals. Perhaps surprisingly, we might want to offer a consumer \$20 for three referrals, but only \$15 for four.

To understand why this happens, consider a simple example in which F is uniform, all consumers have at most three neighbors, the cost of making referrals is zero, and the firm's price and value for advertising are $p = y = \frac{1}{2}$. Figure 1 shows the optimal discounts of $I_0 = 0$, $I_1 = \frac{1}{4}$, and $I_2 = I_3 = \frac{1}{2}$. Suppose the degree distribution is such that the corresponding neighbor purchase probability is $P_{\sigma, \mathbf{w}} = \frac{3}{4}$. We compute the corresponding payment function \mathbf{w} as $\mathbf{w}(1) = \frac{1}{3}$, $\mathbf{w}(2) = \frac{2}{3}$, and $\mathbf{w}(3) = \frac{11}{27} < \frac{2}{3}$. Because of the randomness in neighbor purchase decisions, the discount a consumer receives is a weighted average over values of the payment function \mathbf{w} . A consumer with three neighbors has a higher chance of two successes than a consumer with only two neighbors. In order to equate I_2 and I_3 , we must reduce the benefit from the third successful referral to balance out this effect.

In practice, firms do not employ non-monotonic payments. Not only are these payment

⁵Note that these optimal effective prices are subject to boundary constraints. If p is below the monopoly price, then low-degree consumers may receive no discount at all; sufficiently high degree consumers will always receive a full discount, facing an effective price of zero.

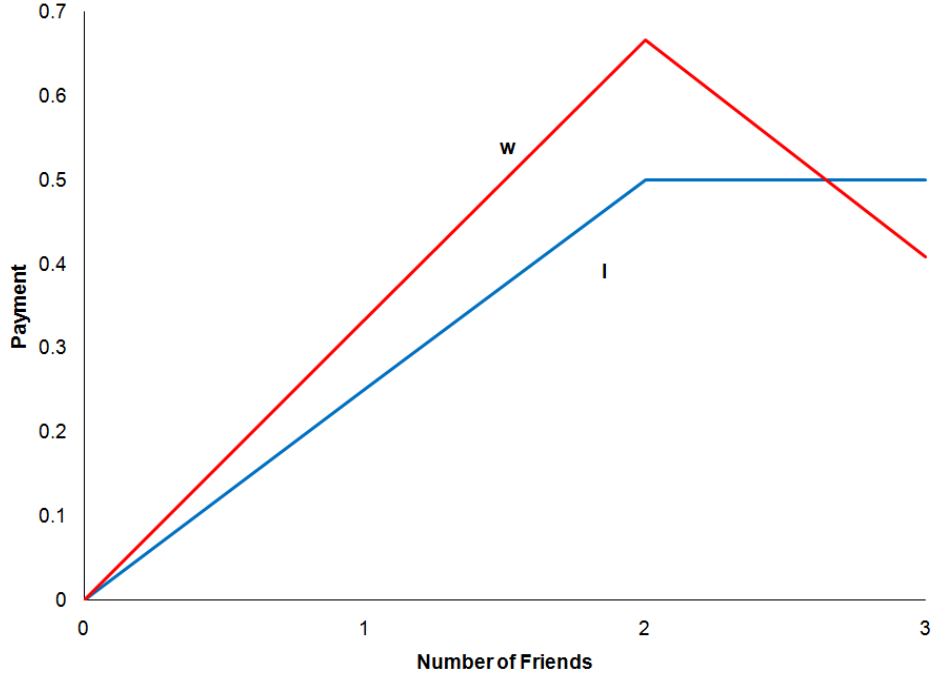


Figure 1: Optimal discounts and non-monotone payments

functions difficult to compute, but real consumers are unlikely to respond optimally. To address this shortcoming, we study approximations based on empirically common referral payment structures. We consider how well linear and threshold payment functions can approximate the optimal scheme, and we look at how the structure of the network and the cost of making referrals affects which type of referral program performs better.

5 Approximate Incentives

A linear payment function \mathbf{w}_L is one that gives the consumer a fixed payment for each successful referral, i.e., $\mathbf{w}_L(N) = aN$ for some $a \geq 0$. Using a linear \mathbf{w}_L , we can implement any sequence of pairs $\{(r_d, I_d)\}$ with $r_d = d$ and $I_d = dI$ for some $I \geq 0$. A threshold payment function \mathbf{w}_T makes a fixed payment b once some threshold number of successful referrals is met. We have $\mathbf{w}_T(N) = b$ if and only if $N \geq \tau$, and $\mathbf{w}_T(N) = 0$ otherwise, for some $b \geq 0$ and $\tau \geq 0$. The set of implementable pairs is less straightforward to describe for threshold

payments as it depends on changes in the probability of crossing a given threshold. In either case, the restriction on the set of pairs $\{(r_d, I_d)\}$ means we lose the ability to decouple the summands in our optimization problem (7): the degree distribution now matters.

Our analysis focuses on worst-case degree distributions. Clearly, some degree distributions entail no approximation loss—for instance, if all consumers have the same number of neighbors. A worst-case analysis helps us understand how robust different types of referral programs are. The optimal profit, and hence the potential approximation loss, scales with the average degree in the network $\mu \equiv \mathbb{E}_{D \sim G}[D]$, holding $y - \delta$ fixed. In order to make meaningful comparisons, we consider worst-case degree distributions with a given average degree, and we shall look at how the worst-case approximation loss scales with the average degree in the network.

Given a degree distribution G , let $\pi^*(G)$ denote the firm’s profit using the optimal referral payments, let $\pi_L(G)$ denote the firm’s profit using the best linear payment function, and let $\pi_T(G)$ denote the firm’s profit using the best threshold payment function. We define the **worst-case loss** from linear payments in a network with average degree μ as

$$L_L(\mu) = \sup_{G: \mathbb{E}_{D \sim G}[D] = \mu} \pi^*(G) - \pi_L(G),$$

and we define the worst-case loss from threshold payments in a network with average degree d as

$$L_T(\mu) = \sup_{G: \mathbb{E}_{D \sim G}[D] = \mu} \pi^*(G) - \pi_T(G).$$

The optimal profit scales linearly in μ as the value of referrals becomes the main benefit the firm obtains. Approximation losses from using linear payments grow much more slowly, scaling with $\sqrt{\mu}$, while those from using threshold incentives grow linearly. This indicates that linear referral payments are typically preferable when μ is large.

Theorem 3. *Approximation losses from using linear payments $L_L(\mu)$ scale according to $\sqrt{\mu}$. That is, there are constants \underline{a} , \bar{a} and a' such that*

$$\underline{a}\sqrt{\mu} \leq L_L(\mu) \leq \bar{a}\sqrt{\mu} \quad \text{for all } \mu \geq a'.$$

If $\delta > 0$, approximation losses from using threshold payments $L_T(\mu)$ scale linearly with μ . That is, there are constants \underline{b} , \bar{b} and b' such that

$$\underline{b}\mu \leq L_T(\mu) \leq \bar{b}\mu \quad \text{for all } \mu \geq b'.$$

Proof. Consider the linear approximation first. To establish the upper bound, note that for sufficiently large average degree μ the optimal policy for all consumers with degree d higher than $\sqrt{\mu}$ is to offer $I_d = p$, yielding profit of $d(y - \delta)$ per consumer. For consumers of degree less than $\sqrt{\mu}$, the optimal profit is no more than $\sqrt{\mu}(y - \delta)$ per consumer since profits per consumer are increasing in degree.

Suppose we set $I_d = dI$ with $I = \frac{p}{\sqrt{\mu}}$ for all d . For the consumers with degree above $\sqrt{\mu}$, the loss due to overpayment is at most $\frac{p}{\sqrt{\mu}}$ per referral. We can bound the loss from overpayments by

$$\mathbb{E}[I_d] = \frac{p\mathbb{E}_{D \sim G}[D]}{\sqrt{\mu}} = p\sqrt{\mu}.$$

Likewise, since the total profit from consumers with degree lower than $\sqrt{\mu}$ is bounded by $\sqrt{\mu}(y - \delta)$, this is a bound on the loss as well because profits from a consumer are never negative. This proves the upper bound.

For the lower bound, consider the degree distribution G_μ that assigns degree $\lfloor \sqrt{\mu} \rfloor$ with probability $\frac{\mu}{2\mu - \lfloor \sqrt{\mu} \rfloor}$ and degree 2μ with probability $\frac{\mu - \lfloor \sqrt{\mu} \rfloor}{2\mu - \lfloor \sqrt{\mu} \rfloor}$. This distribution has expectation μ , and for sufficiently large μ , the optimal policy sets $I_{\lfloor \sqrt{\mu} \rfloor} = I_{2\mu} = p$. Define $I = \frac{I_{\lfloor \sqrt{\mu} \rfloor}}{\lfloor \sqrt{\mu} \rfloor} = \frac{I_{2\mu}}{2\mu}$, and consider two cases.

Suppose first that $I \leq \frac{p}{2\lfloor \sqrt{\mu} \rfloor}$. Then low-degree consumers receive a discount of no more than $\frac{p}{2}$, and lost sales coupled with lost referrals lead to approximation losses of at least $F\left(\frac{p}{2}\right) \lfloor \sqrt{\mu} \rfloor (y - \delta)$ from each low-degree consumer. For large μ , essentially all consumers have low degrees, so this is a lower bound on the approximation loss. Alternatively, suppose $I > \frac{p}{2\lfloor \sqrt{\mu} \rfloor}$. For large μ , essentially all referrals come from high degree consumers, and essentially all of this discount is overpayment. For large μ , the overpayment losses converge to $\frac{(\mu - \lfloor \sqrt{\mu} \rfloor)p}{2\lfloor \sqrt{\mu} \rfloor}$. In both cases, the worst-case loss is at least a constant times $\sqrt{\mu}$, proving the claim.

Now consider the threshold approximation. The upper bound is trivial since optimal profits scale linearly with μ . For the lower bound, consider for each μ the degree distribution that places probability $\frac{1}{2}$ on $\frac{\mu}{2}$ and probability $\frac{1}{2}$ on $\frac{3\mu}{2}$.

Suppose that for some large μ , the threshold τ is larger than $\frac{\mu}{2}$. None of the low-degree consumers make any referrals, which leads to losses of $\frac{\mu}{4}(y - \delta)$. Alternatively, suppose that $\tau \leq \frac{\mu}{2}$. To avoid similar losses, we must incentivize the high-degree consumers to refer at least μ of their neighbors. From Hoeffding's inequality, for large μ , the probability of *not* hitting the threshold number of successes τ after μ referrals is on the order of $e^{-\mu/2}$. If adding one more referral is to increase the probability of hitting τ enough to compensate for the cost δ , the threshold payment must be very large, growing exponentially in μ . This implies overpayments to high-degree consumers that grow exponentially in μ . Hence, for large μ the optimal threshold payment scheme will choose a high threshold τ , yielding linear losses because the low-degree consumers do not refer any neighbors. \square

The proof of Theorem 3 sheds light on why linear and threshold payments perform well or poorly. Linear payments suffer when the degree distribution is heavy-tailed. The discount that a consumer receives is proportional to how many neighbors she has. When most consumers have low degrees, but most referrals come from those with high degrees, linear payments face a difficult tradeoff. Providing the appropriate discount to low-degree consumers entails large overpayments to high-degree ones, while giving an appropriate discount to high-degree consumers entails fewer purchases from low-degree ones. Simple modifications to a linear payment function, like adding a payout cap, do not change the asymptotic result. In the case of linear payments with a payout cap, high-degree consumers simply refrain from referring all of their neighbors, generating comparable losses.

If we can constrain the tail of the degree distribution, the losses from using linear payments scale more slowly. Given a constant $C > 0$ and an exponent $\alpha \in [0, 1)$, consider the collection of all distributions $\mathcal{S}_C^\alpha(\mu)$ with average degree μ such that the tail above the mean satisfies

$$\sum_{d \geq \mu} (d - \mu) g_d \leq C \mu^\alpha.$$

The exponent α is the crucial part of this constraint; with larger α , the permitted length of the tail scales more rapidly with the average degree. We can interpret this as a restriction on how large the mean is conditional on an above average degree. The case $\alpha = 0$ implies the conditional mean is no larger than μ plus a constant, while an $\alpha > 0$ allows the difference between μ and the conditional mean to grow. We define the (C, α) **worst-case loss** from using linear payments as

$$L_C^\alpha(\mu) = \sup_{G \in \mathcal{S}_C^\alpha(\mu)} \pi^*(G) - \pi_L(G).$$

Proposition 3. *Fixing C and α , the (C, α) worst-case loss from using linear payments scales no faster than $\mu^{\alpha/2}$. That is, there exist constants \bar{a} and a' such that*

$$L_C^\alpha(\mu) \leq \bar{a}\mu^{\alpha/2} \quad \text{for all } \mu \geq a'.$$

Proof. Set $I = \frac{1}{2\mu^{1-\alpha/2}}$. For sufficiently large μ , the optimal policy gives $I = p$ to all consumers with degree higher than $\frac{\mu^{1-\alpha/2}}{2}$, and the loss from over payments to high degree consumers is no more than $\frac{\mu}{2\mu^{1-\alpha/2}} = \frac{\mu^{\alpha/2}}{2}$. Given our constraint on the tail of the degree distribution, we can find a constant C' such that the probability that a consumer has degree less than $\frac{\mu^{1-\alpha/2}}{2}$ is no more than $\frac{C'}{\mu^{1-\alpha}}$. Consequently, the loss from consumers with degree less than $\frac{\mu^{1-\alpha/2}}{2}$ is at most proportional to $\frac{\mu^{1-\alpha/2}}{2} \frac{C'}{\mu^{1-\alpha}} = \frac{C'\mu^{\alpha/2}}{2}$, completing the proof. \square

Though threshold payments have worse asymptotic properties, the conditions that hamper their performance are distinct. Having just one payment limits the flexibility of these schemes to compensate the marginal cost of making referrals. If these costs are significant, then a threshold payment can only target consumers within a narrow range of degrees. Those with lower degrees cannot benefit from the threshold payment since they will never hit the threshold. Those with higher degrees still receive an appropriate discount, but they refrain from referring all of their neighbors because the marginal impact on the probability of hitting the threshold is too small to compensate for the additional cost.

In networks with relatively low average degrees, these different strengths and weaknesses can make either linear or threshold payments preferable. To see this, consider an example comparing the performance of linear and threshold payments across different networks. For

these calculations, we assume consumer valuations are uniform, the price of the product is $p = 0.5$, the cost of making referrals is zero, and the value of referrals to the firm is $y = 0.08$. We compute optimal linear and threshold payments for three different degree distributions: a Poisson distribution, a geometric distribution, and a power law distribution with exponent 2.1. In each case, the average degree in the network is approximately 3.6. Figure 2 reports the percentage of the optimal profit we obtain in each case.

	Poisson	Geometric	Power Law
Linear Payment	100%	99.1%	90.8%
Threshold Payment	99.4%	97.9%	98.5%

Figure 2: Percentage of Optimal Profit

Performance is quite good across the board; having a referral program, even a very simple one, is far better than having none. We can see a clear decline in the performance of linear payments as the tail of the degree distribution becomes fatter. Moving from a Poisson to a geometric distribution costs one percent of profits, and moving to the much fatter-tailed power law distribution costs an additional eight percent. Threshold payments are more consistent across the distributions. For the Poisson and the geometric distribution, threshold payments perform slightly but noticeably worse than linear ones. For the power law distribution, threshold payments offer a significant improvement. Note the absence of referral costs is important to this comparison. If we increase δ while holding the net referral value $y - \delta$ fixed, the performance of linear payments will remain unchanged, but threshold payments will suffer.

Linear and threshold payments have complementary strengths. Linear payments can effectively compensate the cost of making referrals, while threshold payments can provide a discount to high-degree consumers that remains bounded. A combination of these two can lead to much better performance. Consider the class of linear payment functions with a

“bonus” for crossing some threshold. These are payment functions \mathbf{w}_B we can express as

$$\mathbf{w}_B(N) = \mathbf{w}_L(N) + \mathbf{w}_T(N),$$

where \mathbf{w}_L is linear and \mathbf{w}_T is a threshold payment function. Let $L_B(\mu)$ denote the worst-case loss using such payment functions when the average degree in the network is μ . With this slight increase in complexity, we can achieve a constant bound on losses.

Theorem 4. *The worst-case loss using linear payments with a bonus is bounded by a constant. There exists $C > 0$ such that*

$$L_B(\mu) \leq C \quad \text{for all } \mu.$$

Proof. We prove a stronger result, showing that a large class of policies achieves constant loss. Choose the linear component \mathbf{w}_L to exactly compensate the social cost δ per referral. This reduces our problem to showing that a threshold payment achieves constant loss when the social cost of referrals is zero.

In the specification of the threshold payment function \mathbf{w}_T , choose any $c > p$ and any finite K . Since $P_{\sigma, \mathbf{w}} \geq 1 - F(p) > 0$, there is some threshold degree \bar{d} such that for any purchasing consumer with $d \geq \bar{d}$, the probability that at least K neighbors purchase in response to a referral is at least $\frac{p}{c}$. These consumers receive a full discount and overpayment bounded by $c - p$. Losses from consumers with a degree lower than \bar{d} , or for whom the optimal policy does not offer a full discount, are bounded by a constant because there are finitely many such degrees. Losses from higher degree consumers are bounded by $c - p$. \square

To avoid significant losses from the highest degree consumers, we need to offer a full discount and we need to compensate for each referral. Choosing a linear component that exactly compensates for the social cost δ ensures that any purchasing consumer makes all possible referrals, and we can adjust the threshold payment to provide a sufficiently large discount for the high-degree consumers. The size of the threshold payment gives us a fixed bound on any over payments, yielding the constant loss bound in Theorem 4.

An implication of Theorem 4 is that, in the absence of any cost for making referrals, threshold payments alone can achieve constant loss.

Corollary 1. *If $\delta = 0$, approximation losses from using threshold payments $L_T(\mu)$ are bounded by a constant.*

Together with our earlier findings, this suggests that threshold payments can outperform linear ones in some settings. Whenever we have that

(a) The degree distribution has a fat tail, so linear payments perform relatively poorly, and

(b) The marginal cost of referrals is low, so threshold payments perform relatively well,

then threshold payments may be the best option among the simplest payment schemes. Notably, conditions (a) and (b) are characteristic of many social media settings: the degree distribution in many social networks follows a power law [e.g. Ugander et al., 2012], and referrals can often be made en masse using a single post.

6 Extensions

6.1 Differentiating Conversions from Non-Conversions

Though there is certainly value to a company from referrals beyond any immediate conversions, it makes sense that a firm would place a higher value on referrals that convert than referrals that do not. We can extend our model to account for this, supposing that the firm values unconverted referrals at \underline{y} and converted referrals at \bar{y} with $\bar{y} > \underline{y} > \delta$. The expected value of a referral is then endogeneously determined in equilibrium as $y = P_{\sigma, \mathbf{w}} \bar{y} + (1 - P_{\sigma, \mathbf{w}}) \underline{y}$.

Although this adds an extra dimension to the firm's optimization problem, the essence of our results remains unchanged. The firm's problem in equation (7) no longer decouples along player degrees since y depends on the strategies of all players, but we can still solve a relaxed version of the problem that only imposes individual rationality constraints. The solution of the relaxed problem will still be incentive compatible, so we can find a payment function to implement the optimal sequence of pairs $\{(r_d, I_d)\}$. Knowing that $\delta < \underline{y} \leq y \leq \bar{y}$ is enough to reproduce the asymptotic bounds of Section 5 using the same arguments.

6.2 Optimal Pricing

If the firm jointly optimizes the product price and the referral program, how should the price be set? If there are no Veblen effects and we can use the optimal payment function, then the answer is obvious: choose the standard monopoly price. For consumers with no friends to refer, the monopoly price maximizes profits, and the optimal payment function allows us to give the best possible sequence of discounts from that reference point. If all consumers have at least one neighbor, then the optimal price is non-unique: anything higher than the monopoly price is optimal. This is because we can implement an arbitrary non-decreasing sequence of discounts, so the effective prices we charge consumers will not change with the price p as long as it is above the monopoly price.

If we are using an approximate payment function, the optimal price is less clear. Since the payment function depends on fine details of the degree distribution, the optimal price will as well. At least in the case of linear incentives, we can make a general claim about the optimal price: the standard monopoly price serves as an upper bound.

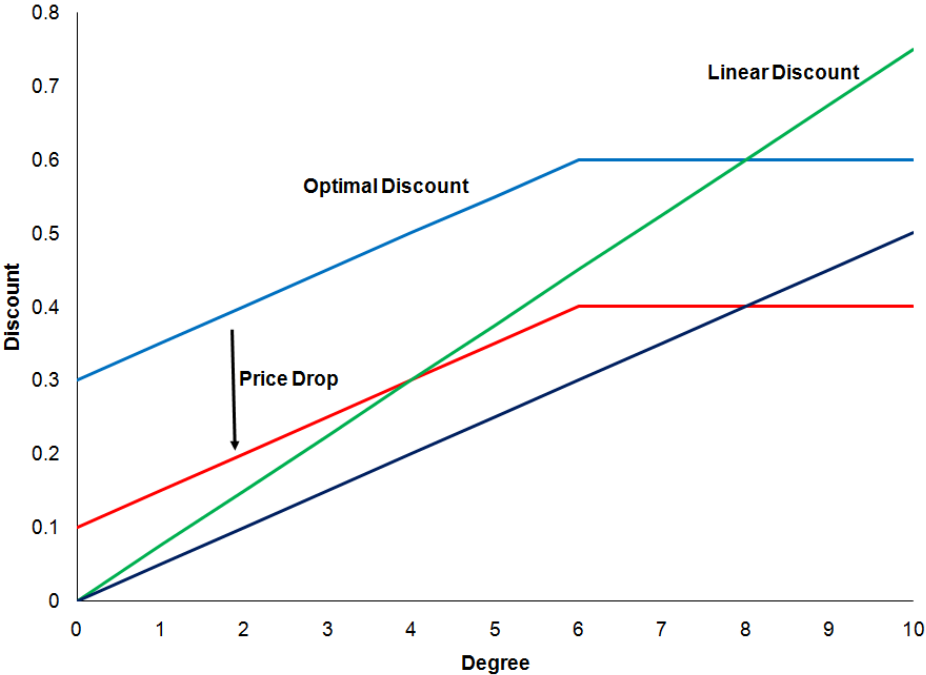


Figure 3: If p is above the monopoly price, reducing p results in a better approximation.

Proposition 4. *Suppose virtual valuations $v - \frac{1-F(v)}{f(v)}$ are increasing and (weakly) convex. Then if we are using a linear payment function \mathbf{w} , the optimal price is no larger than the standard monopoly price.*

Proof. First, suppose the valuation distribution is uniform. Figure 3 shows the optimal discount as a function of degree as well as a linear approximation. Decreasing the price p shifts the optimal discount curve downwards, and the curve hits the point $(0, 0)$ precisely when p is the standard monopoly price. Note that for p equal or greater than the monopoly price, any linear approximation intersects the curve at exactly one point. If we shift the curve downward we can always obtain a strictly better approximation by flattening the line to maintain the same point of intersection: the resulting line is closer to the optimal incentive at every point. This same graphical argument applies whenever the optimal discount curve is weakly concave. Since we have assumed virtual valuations are convex, Eq. (10) implies that the optimal discount curve is concave, yielding the result. \square

6.3 Two Way Incentives

Many referral programs make use of two-sided incentives: in addition to giving the referring customer a reward, the referred friend receives a discount. Why should the firm offer such incentives? In our framework, there is little reason to include this feature since we already have as much flexibility as we could want to adjust the expected referral reward. The only effect would be to reduce revenues from referred individuals, which would imply the firm should not offer this type of incentive. However, given the prevalence of such programs, this answer is unsatisfying.

One possible explanation is that the additional incentive may reduce social costs. Part of the cost of making referrals is psychological, stemming from a desire not to annoy one's friends or appear self-serving. If making a referral confers a benefit to the friend receiving it, these psychological costs may shrink. We could extend our model to make the social cost δ a function of the discount provided to the friend receiving the referral. The firm would then jointly optimize the referral reward function and the net referral value by manipulating the

social cost. Once the social cost is fixed by choosing the discount for referred friends, the problem reduces to the one we study, so our principal findings remain unchanged.

7 Final Remarks

Start-up companies increasingly rely on referrals to launch their businesses, and many of the great success stories in recent memory have been driven by customer referral programs. The rise of social media makes it easier for companies to track referrals and provide associated payments, greatly reducing the costs of implementing these programs relative to traditional advertising. Social media also makes it easier for consumers to refer their friends; e-mailing a link or making a Facebook post requires far less effort than bringing up a product in conversation. These features will ensure a continuing role for referral programs in firms' advertising mix.

Our decomposition of referral payments highlights two distinct considerations when we design a referral program: we need to compensate consumers for making referrals, and we need to discount the product for those who can bring in many referrals. Since we never wish to pay the consumer for buying the product, the latter piece introduces a kink in the optimal discount as a function of how many neighbors a consumer has. With no restrictions on our referral payment function, this kink can lead to unwieldy optimal payments that are non-monotonic in the number of successful referrals.

The same decomposition can guide our thinking when we design simple payment schemes to approximate the optimal one. In our model, linear payments can effectively compensate consumers for making referrals, and threshold payments can provide a discount to high-degree consumers while limiting over payments. As a result, each type of referral program has different strengths and weaknesses depending on the structure of the social network and the significance of referral costs. Linear payments are more flexible in general and will typically out perform threshold payments. However, the combination of a fat-tailed degree distribution with low marginal referral costs, which is characteristic of many social media settings, favors threshold payments.

If we can implement a slightly more complex payment scheme, linear payments with a bonus may offer a substantially better option. These functions combine features of both linear and threshold payments. We have two degrees of freedom, one that we can tailor to compensate consumers for making referrals and one that we can use to appropriately discount the product for high-degree consumers. Regardless of the underlying network structure or the cost of referrals, this type of referral program can offer a good approximation to the optimal profit.

References

- Sinan Aral and Dylan Walker. Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks. *Management Science*, 57(9):1623–1639, 2011.
- Sinan Aral, Lev Muchnik, and Arun Sundararajan. Engineering Social Contagions: Optimal Network Seeding in the Presence of Homophily. *Network Science*, 1(2):125–153, 2013.
- Coralio Ballester, Antoni Calvó-Armengol, and Yves Zenou. Who’s who in networks. wanted: the key player. *Econometrica*, 74(5):1403–1417, 2006.
- Eyal Biyalogorsky, Eitan Gerstner, and Barak Libai. Customer Referral Management: Optimal Reward Programs. *Marketing Science*, 20(1):82–95, Winter 2001.
- Francis Bloch and Nicolas Quérou. Pricing in Social Networks. *Games and Economic Behavior*, 80:243–261, 2013.
- Moira Burke and Robert Kraut. Growing Closer on Facebook: Changes in Tie Strength Through Social Network Site Use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI2014)*, pages 4187–4196, 2014.
- Francis A. Buttle. Word of Mouth: Understanding and Managing Referral Marketing. *Journal of Strategic Marketing*, 6(3):241–254, 1998.

- Arthur Campbell. Word-of-Mouth Communication and Percolation in Social Networks. *American Economic Review*, 103(6):2466–2498, 2013.
- James Campbell. Localized Price Promotions as a Quality Signal in a Publicly Observable Network. *Quantitative Market and Economics*, 13(1):27–57, 2015.
- Ozan Candogan, Kostas Bimpikis, and Asuman Ozdaglar. Optimal Pricing in Networks with Externalities. *Operations Research*, 60(4):883–905, 2012.
- Kalyan Chatterjee and Bhaskar Dutta. Credibility and Strategic Learning in Networks. Working Paper, 2014.
- Shu-Chuan Chu and Yoojung Kim. Determinants of Consumer Engagement in Electronic Word-of-Mouth (eWOM) in Social Network Sites. *International Journal of Advertising*, 30(1):47–75, 2011.
- Maxime Cohen and Pavithra Harsha. Designing Price Incentives in a Network with Social Interactions. Working Paper, 2015.
- Andrea Galeotti, Sanjeev Goyal, Matthew Jackson, Fernando Vega-Redondo, and Leat Yariv. Network Games. *Review of Economic Studies*, 77(1):218–244, 2010.
- David Godes and Dina Mayzlin. Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science*, 23(4):545–560, 2004.
- Jacob Goldenberg, Barak Libai, and Eitan Muller. Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. *Marketing Letters*, 12(3):211–223, 2001.
- Drew Houston. Dropbox – Startup Lessons Learned. *Presentation slides available at <http://www.slideshare.net/gueste94e4c/dropbox-startup-lessons-learned-3836587>*, 2010.
- David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the Spread of Influence through a Social Network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 137–146. ACM, 2003.

- V. Kumar, Andrew Petersen, and Robert Leone. How Valuable Is Word of Mouth? *Harvard Business Review*, 85(10):139–144, 2007.
- Jure Leskovec, Lada A Adamic, and Bernardo A Huberman. The Dynamics of Viral Marketing. *ACM Transactions on the Web*, 1(1):5, 2007.
- Barak Libai, Eyal Biyalogorsky, and Eitan Gerstner. Setting Referral Fees in Affiliate Marketing. *Journal of Service Research*, 5(4):303–315, 2003.
- Markus Mobius, Paul Niehaus, and Tanya Rosenblat. Social Networks and Consumer Demand. Working Paper, 2011.
- Roger Myerson. Optimal Auction Design. *Mathematics of Operations Research*, 6(1):58–73, 1981.
- Arun Sundararajan. Local Network Effects and Complex Network Structure. *The B. E. Journal of Theoretical Economics*, 7(1), 2007.
- Johan Ugander, Brian Karrer, Lars Backstrom, and Cameron Marlow. The Anatomy of the Facebook Social Graph. Working Paper, 2012.