CONTRACTING WITH WORD-OF-MOUTH MANAGEMENT

By

Yuichiro Kamada and Aniko Öry

July 2016

COWLES FOUNDATION DISCUSSION PAPER NO. 2048

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Contracting with Word-of-Mouth Management

Yuichiro Kamada and Aniko Öry*

July 12, 2016

Abstract

We incorporate word of mouth (WoM) in a classic Maskin-Riley contracting problem, allowing for referral rewards to senders of WoM. Current customers' incentives to engage in WoM can affect the contracting problem of a firm in the presence of positive externalities of users. We fully characterize the optimal contract scheme and provide other comparative statics. In particular, we show that offering a free contract is optimal only if the fraction of premium users in the population is small. The reason is that by offering a free product, the firm can incentivize senders to talk by increasing expected externalities that they receive and this can (partly) substitute for paying referral rewards only if there are few premium customers. This result is consistent with the observation that companies that successfully offer freemium contracts oftentimes have a high percentage of free users.

*Kamada: Haas School of Business, University of California, Berkeley, Berkeley, CA 94720, e-mail: y.cam.24@gmail.com; Öry: School of Management, Yale University, New Haven, CT 06511, e-mail: aniko.oery@yale.edu. We are grateful to Juan Escobar, Johannes Hörner, Fuhito Kojima, Vineet Kumar, Dina Mayzlin, Takeshi Murooka, Motty Perry, Klaus Schmidt, Jiwoong Shin, Philipp Strack, Steve Tadelis, Juuso Välimäki, Miguel Villas-Boas, Alex Wolitzky, Jidong Zhou, and seminar participants at the 13th Summer Institute in Competitive Strategy, Osaka University, the University of Munich (LMU), Wharton, and Yokohama National University for helpful comments. Jovian Chen, Don Hampton, Yi (Michelle) Lu, Valeree Simon, and Tawney Warren provided excellent research assistance.

"Cost per acquisition: \$233-\$388. For a \$99 product. Fail."

—Drew Houston, founder of Dropbox

1 Introduction

When Dropbox went public in 2009 without offering a referral reward and hiding its free trial option, costs per acquisition were more than 200 dollars for a 99 dollar product. In April 2010, Dropbox completely changed its strategy by starting its referral program, increasing visibility of its free 2 GB option, and introducing a sharing option. All in all, this led to 2.8 million direct referral invites within 30 days (Houston (2010)). Companies such as Uber, Amtrak and many others, too, have offered various referral programs to date. For example, Uber doubled referral credits for the new year in 2014 which they announced on UBER Newsroom (2014). Despite the prevalence of incentive schemes to encourage WoM, the theoretical literature on WoM and contracting, besides a few exceptions that we will discuss later, has thus far largely ignored the incentives to talk and has instead focused on the mechanical processes that model the spread of information.The objective of this paper is to investigate the question of how firms should encourage customers to engage in WoM and how the optimal contract scheme, in particular the use of free contracts and referral rewards, interact with network externalities in this context.

In order to find the optimal incentive scheme, it is crucial to understand why people talk.[1] Senders of information (existing customers) face a tradeoff generated by three actors in the market — themselves, the receivers (potential new customers), and the firm. On the one hand, there are many reasons why talking is costly: senders incur opportunity costs of talking (Lee et al., 2013), and/or they may feel psychological barriers. On the other hand, senders can benefit from advertising the product they use: they can receive referral rewards from the firm, while receivers generate *positive externalities*. Such an externality

[1]Berger (2014) surveys the behavioral studies that examine why people talk, and argues that people's motivation is self-serving, even without their awareness. See also Berger and Schwartz (2011) for a field experiment on the psychological drivers of WoM.

can be a real value of social usage or psychological benefit from having convinced a friend to use the same product (Campbell et al., 2015). The sender may also benefit from the continuation value in a repeated relationship with the receiver. Whether the sender can enjoy such externalities depends on whether the receiver uses the product, and the firm can affect the likelihood of usage by fine-tuning the menu of contracts offered to the receivers. Specifically, the firm can increase the expected number of receivers who use the product, and thus the *expected* externalities, by offering free contracts. This is because receivers who would not have purchased the product otherwise will then use it. All in all, each sender wants to talk if and only if

$$\underbrace{\text{Cost of talking}}_{\text{Internal to the sender}} \leq \underbrace{\text{Referral rewards}}_{\text{Provided by the firm}} + \underbrace{\text{Expected externalities}}_{\text{Generated by the receivers}}.$$

In this paper, we aim to understand the implication of this tradeoff on the firm's optimal contracting scheme. For that purpose, we enrich a classic contracting problem as in Maskin and Riley (1984) by allowing the number of customers to depend on the referral decision by the senders of information, who face the aforementioned tradeoff. In the simplest setting in which cost of talking is homogeneous across agents, we completely characterize the optimal scheme. It exhibits a rich pattern of the use of referral rewards and free products, depending on the parameters in the model. Roughly speaking, the model predicts that referral rewards are used if externalities are low, and free products are used only if the fraction of "premium users" is low. We also show that, for referral rewards to be used in conjunction with free contracts, the externalities cannot be too low, although they cannot be too high either. Such predictions are consistent with observed contracts in reality: Skype (a telecommunication application with about only 8% of paying customers) uses only free products but not referral rewards,[2] Uber and Amtrak (ground transportation services) use only referral rewards, and Dropbox (a cloud storage and file synchronization service with a sharing option and about

[2]Another product that falls into this category would be LinkedIn (a social networking service with less than 1% of paying customers (Woirhaye, 2011)

only 4% of paying customers (*Economist*, 2012) uses both free products and referral rewards.

The key intuition for these results is the following. If externalities are high, the firm does not need to provide additional incentives for talking by giving away referral rewards. This is why rewards are used only when externalities are small. The reason to use free contracts is to boost up the expected externalities that the sender receives. The "jump" of the expected externalities is large (and thus effective) only when the fraction of users who would otherwise not use the product is high and externalities are not too small. This is why free products are used only when the fraction of high types is small.

The exact tradeoff is more complicated than this. One such complication pertains to the cost of free products. Note that the discussion so far only describes the magnitude of the benefit of offering a free product, while ignoring its cost. There are two reasons that such a strategy is costly. First, the firm incurs a production cost of the free product (which is low for products such as Skype and Dropbox). Second, it might have to pay an information rent to high-valuation buyers. This total cost of offering free contracts plays a key role in fully characterizing the optimal incentive scheme. Another complication is that there is non-monotonicity of the use of rewards with respect to the size of externalities. That is, it is possible that the optimal reward changes from positive to zero when externalities are lowered because free contracts can "substitute" rewards. We formalize what we mean by substitution, and explain how the two strategies (rewards and free contracts) interact in characterizing the optimal scheme. This effect results from the aforementioned incentive constraint of the sender.

To the best of our knowledge, the present paper is the first that takes into account externalities generated by communication in the context of WoM. The existence of such externalities rationalizes how companies that use free contracts such as Skype and Dropbox were able to create a buzz for their product, while for markets with lower externalities or a higher fraction of high valuation buyers, such as providers of transportation (e.g., Uber or Amtrak), a classic reward program is the optimal strategy. Importantly, existence of externalities in our model explains referral programs and free contracts in a unified framework. It allows us to understand how the two seemingly unrelated popular marketing techniques interact with each other as part of the optimal marketing mix for a firm.

The paper is structured as follows. Section 2 introduces the model, analyzes the benchmark case in which there is no cost of talking for senders, and demonstrates some basic properties that the optimal menu of contracts and referral reward schemes must satisfy. The main analysis presented in Section 3 is concerned with the case in which the cost of talking is homogenous across senders. We completely characterize the optimal menu of contracts and referral reward scheme, and conduct comparative statics. Section 4 discusses various extensions, robustness checks, and welfare considerations. Section 5 concludes. Proofs are deferred to the Appendix. The detailed analysis of the model with heterogeneous cost of talking can be found in the Online Appendix.

1.1 Related Literature

This paper contributes to the literature on WoM management. To the best of our knowledge there are only two recent papers that are concerned with the question of how the firm can affect the strategic communication behavior of their customers.[3] Our paper is the closest to Biyalogorsky et al. (2001) who compare the benefits of price reduction and referral programs in the presence of WoM. In their model, a reduced price offered to the sender of WoM is beneficial because it makes the sender "delighted" and thereby encourages him to talk. Depending on the delight threshold, the seller should use one of the two strategies or both. In contrast, our focus is on WoM in the presence of positive externalities of talking and our model accommodates menus of contracts. In Campbell et al. (2015), senders talk in order to affect how they are perceived by the receiver of the information. The perception is better if the information is more exclusive. Thus, a firm can improve overall awareness of the product by restricting access to information (i.e., by advertising less). One could interpret

[3]See also Godes et al. (2005) for a survey of the literature.

the positive externalities in our model also as a reduced form of a "self-enhancement motive" as in their model. Although we discuss advertising in Section 4.5, we focus on the relative effectiveness of free contracts and referral rewards instead of advertising.

Most of the other theoretical literature on WoM has focused on mechanical processes of communication in networks. This literature mostly focuses on how characteristics of the social network affect a firm's optimal advertising and pricing strategy. Campbell (2012) analyzes the interaction of advertising and pricing.[4] Galeotti (2010) is concerned with optimal pricing when agents without information search for those with information. Galeotti and Goyal (2009) show that advertising can become more effective in the presence of WoM (i.e., WoM and advertising are complements) or it can be less effective because WoM attracts more people than advertising can (i.e., WoM and advertising are substitutes). All of these papers consider information transmission processes in which once a link is formed between two agents, they automatically share information.

Costly communication has been studied in the context of working in teams where moral hazard problems are present between the sender and receiver, as introduced by Dewatripont and Tirole (2005). Dewatripont (2006), for example, applies their model to study firms as communication networks. Instead, our model does not involve moral hazard but a screening problem, and externalities (which are absent in Dewatripont (2006)) play a key role in formulating the optimal contracting scheme.

There is also a literature on contracting models in the presence of network effects. Besides the critical difference that our focus is on how the firm can optimally affect WoM, there is a subtle difference in the optimal contracts. Csorba (2008) analyzes such a contracting model with network effects, in which the more the other buyers use the product, the higher the utility from using the product is.[5] He shows that an optimal contract scheme introduces a distortion at the top because a reduction of the quantity offered to low types should decrease

[4] On the empirical side, Godes and Mayzlin (2009) analyze the roles of loyal customers and opinion leaders in the context of word-of-mouth.

[5] See Segal (2003) for a seminal work on this literature. See also Hahn (2003).

the value of the product to high types. Unlike in his model, we have no distortions at the top in the optimal contract scheme. The reason is that receivers do not receive externalities from each other, and that we consider quantity-independent externalities rather than assuming that the total quantity consumed generates externalities. We discuss the implications of quantity-dependent externalities in Section 4.3 and show robustness of our results. We do not consider the case of externalities between receivers themselves given our focus on the sender's incentives to talk. Introducing such a feature would not change the qualitative results on the optimal incentive schemes to encourage WoM.

While the focus of this paper is not to add another rationale for freemium strategies, it is important to note the connection to the literature on "freemium" strategies. Lee et al. (2015) empirically analyze the trade-off between growth and monetization under the use of freemium strategies. In their paper, the value of a free customer is determined by the option value of switching from a free contract to a premium contract and by the value of referring a new customer. Our paper shows that there is potentially another value of free contracts, namely the value of encouraging referrals which has been ignored in previous works.

A recent working paper by Shi et al. (2015) considers a static model of product line design without WoM when free users generate positive externalities on all premium users. When the firm can manipulate the amount of externalities enjoyed by customers conditional on the user type, freemium contracts can arise as an optimal strategy. In contrast, in our model, there is no manipulation of the size of externalities and the price of the low-type contracts must be zero because the surplus from selling to the low types is negative. Even so, the monopolist sells contracts to the low types because free contracts encourage WoM which attracts premium users.

Besides these mechanisms for free contracts, Shapiro and Varian (1998) has identified various other reasons: (i) free contracts may be useful in penetration of customers or information transmission about the quality of the product to them, which can induce their upgrade, (ii) the firm may hope that the free users will refer someone who will end up using the premium version,[6] (iii) free products attract attention of customers and prevent them from purchasing the competitors' products. None of these reasons pertains to the senders' incentives. Instead, our focus (with regards to free contracts) is on how free contracts help firms to manage senders' incentives. Thus, instead of convoluting our model with these other aspects of free contracts, we aim to *isolate* the effect of the tradeoff that the senders of information face. Similarly, we do not intend to create a "complete" model that incorporates all conceivable features that are relevant for firms' decision making. Instead, the goal of this paper is to understand how the incentives for WoM can be managed. Our simplification allows us to isolate the factors pertaining to the encouragement of WoM and to examine the tradeoffs involved.

2 Model and Preliminaries

2.1 Model

Basics. We consider a monopolist producing a single product at constant marginal cost $c > 0$. *Senders* (male) $\{1, \ldots, N\}$ can inform *receivers*, (female) $\{1, \ldots, N\}$ about the existence of the product. The monopolist's goal is to maximize the expected profit generated by receivers by offering them a menu of contracts (as in Maskin and Riley (1984)) and, in addition, offering a referral scheme to senders.

Receivers' preferences. Each receiver privately observes her type $\theta \in \{L, H\}$ that determines her valuation of the product. It is drawn independently such that a receiver is of type H with probability $\alpha \in (0, 1)$ and of type L otherwise. A type-θ receiver is associated with a valuation function $v_\theta : \mathbb{R}_+ \to \mathbb{R}$ that assigns to each quantity (or quality) q her valuation $v_\theta(q)$. Over the strictly positive domain, i.e., $q \in (0, \infty)$, we assume that v_θ is continuously differentiable, strictly increasing, strictly concave, $v_H(q) > v_L(q)$, $v'_H(q) > v'_L(q)$ for all q and $\lim_{q\to\infty} v'_H(q) < c$. We assume that $v_H(0) = v_L(0) = 0$, which can be interpreted

[6] A recent working paper by Ajorlou et al. (2015) builds a social-network model that highlights this effect.

as the utility of the outside option of not using the product at all. We make the following additional assumptions:

Assumptions. 1. **(Minimum quantity for low types)** $\exists \underline{q} > 0$ such that $v_L(\underline{q}) = 0$.

2. **(No gains from trade with low types)** $v_L'(q) < c$ for all $q \geq \underline{q}$.

3. **(Gains from trade with high types)** There exists a $q > 0$ such that $v_H(q) > q \cdot c$.

The first assumption can be interpreted as low types incurring some fixed installation cost of the product, and the low valuation buyer only wanting to start using the product if a minimum quantity of $\underline{q} > 0$ is consumed.[7] This assumption together with the normalization that $v_L(0) = 0$ implies that the function v_L is necessarily discontinuous at $q = 0$.[8] The second assumption captures that there are some consumers who would never use the product if they were not needed to incentivize WoM. Without the third assumption, the monopolist would not be able to earn positive profits, so the problem becomes trivial.

Senders' preferences and WoM technology. First, each sender i observes the monopolist's choice of menu of contracts and referral scheme (specified below). Then, he privately observes his cost of talking ξ_i, drawn from an independent and identical distribution with a cumulative distribution function $G : \mathbb{R}_+ \to [0, 1]$. We assume that G has at most finitely many jumps.[9] Each sender i then decides whether to inform receiver i or not.[10] We denote sender i's action by $a_i \in \{\text{Refer}, \text{Not}\}$, where $a_i = \text{Refer}$ if sender i refers receiver i and $a_i = \text{Not}$ otherwise. If (and only if) receiver i learns about the product, she decides whether to purchase a contract or not, and whether to consume the product or not upon purchasing. If receiver i consumes a positive quantity, sender i receives *externalities* $r \geq 0$.[11]

[7]Note that this does not preclude the possibility of positive fixed installation costs for high types.

[8]Recall that continuous differentiability of v_L is assumed only on the strictly positive domain.

[9]The function G has a jump, for example, when all senders share the same cost, which is the case we analyze in Section 3.

[10]Section 4.6 examines the case in which multiple senders talk to a single receiver.

[11]While we set up the problem such that the referred customer does not receive r for notational simplicity, assuming that they do would not change the essence of our analysis. We discuss this point in Section 4.2.

Monopolist's problem. As in Maskin and Riley (1984), the monopolist offers a menu of contracts given by $((p_L, q_L), (p_H, q_H)) \in (\mathbb{R} \times \mathbb{R}_+)^2$ to receivers, where q_θ is the quantity type θ is supposed to buy at a price p_θ. Furthermore, she offers a reward scheme $\mathbf{R} : \{L, H\} \to \mathbb{R}_+$ such that a sender receives $\mathbf{R}(\theta)$ if he has referred a receiver who purchases the θ-contract. Rewards are assumed to be nonnegative because otherwise senders would be able to secretly invite new customers. We assume that the monopolist only receives revenue from new customers who do not know about the product unless a sender talks to them. In order to exclusively focus on the senders' incentive to talk, we assume that the monopolist receives no revenue from senders. Thus, the monopolist solves

$$\begin{aligned} \max_{p_L,\, p_H \in \mathbb{R},\, q_L, q_H \geq 0,\, \mathbf{R} \in \mathbb{R}_+^{\{L,H\}}} \quad & \mathbb{E}_G\Big[\sum_{i=1}^{N} \mathbf{1}_{\{a_i = \text{Refer}\}} \cdot \Big(\underbrace{\alpha \cdot (p_H - q_H \cdot c)) + (1-\alpha) \cdot (p_L - q_L \cdot c)}_{\text{total average profit per referred receiver}} \\ & - (\alpha \mathbf{R}(H) + (1-\alpha)\mathbf{R}(L))\Big)\Big] \end{aligned} \tag{1}$$

subject to the incentive compatibility and participation constraints given by

$$\left.\begin{array}{rcll} \max\{v_H(q_H), 0\} - p_H & \geq & \max\{v_H(q_L), 0\} - p_L & \text{(H-type's IC)} \\ \max\{v_L(q_L), 0\} - p_L & \geq & \max\{v_L(q_H), 0\} - p_H & \text{(L-type's IC)} \\ \max\{v_H(q_H), 0\} - p_H & \geq & 0 & \text{(H-type's PC)} \\ \max\{v_L(q_L), 0\} - p_L & \geq & 0 & \text{(L-type's PC)} \end{array}\right\} \tag{2}$$

and for all i, $a_i = \text{Refer}$ if and only if

$$\xi \leq r\left(\alpha + (1-\alpha) \cdot \mathbf{1}_{\{q_L > 0, v_L(q_L) \geq 0\}}\right) + (\alpha \mathbf{R}(H) + (1-\alpha)\mathbf{R}(L)) \quad \text{(Senders' IC)}$$

Let Π^* denote the value of this problem. The monopolist chooses contracts given by quantities and prices, while *managing WoM*. The management of WoM appears as the senders' incentive compatibility (IC) constraint. The quantity sold to L-type receivers q_L affects WoM by controlling the *expected externalities* given by $r\left(\alpha + (1-\alpha) \cdot \mathbf{1}_{\{q_L > 0, v_L(q_L) \geq 0\}}\right)$. The senders' optimal decision determines the value of the indicator function in the objective function and thereby controls the number of informed receivers.

Let us explain a few assumptions implicit in this formulation. First, as standard in contract theory, we assume tie-breaking conditions for senders and receivers that are most favorable for the monopolist. Senders who are indifferent between referring and not will refer, and receivers that are indifferent between buying and not buying always buy. Second, we assume that if the buyer purchases a contract (p, q) such that $v_\theta(q) < 0$, then the monopolist cannot "force" the receiver to consume even if she pays the buyer a negative price. Thus, a type-θ receiver who purchases such a contract enjoys utility $\max\{v_\theta(q), 0\}$.

2.2 Benchmark with free WoM

We first consider a benchmark case where $G(\xi) = 1$ for all $\xi \geq 0$, i.e., WoM is costless and customers are automatically informed about the product. Then, the monopolist simply solves the classic problem as in Maskin and Riley (1984):

$$\Pi^{\text{classic}} \equiv \max_{p_H,\, p_L \in \mathbb{R}\; q_H, q_L \geq 0} \qquad \alpha \cdot (p_H - q_H \cdot c) + (1-\alpha) \cdot (p_L - q_L \cdot c)$$

subject to the constraints (2). It is always optimal for the seller not to sell to L-type buyers such that $q_L^* = 0$ and the optimal quantity q_H^* sold to H-type buyers satisfies $v_H'(q_H^*) = c$. Assumption 3, strict concavity, continuous differentiability of v_H and $\lim_{q\to\infty} v_H'(q) < c$ ensure that there is a unique such q_H^*. The price for high types is given by $p_H^* = v_H(q_H^*)$ and the maximal static profit is $\Pi^{\text{classic}} = \alpha \cdot (p_H^* - q_H^* \cdot c)$. All in all, we can summarize our findings as follows:

$$v_H'(q_H^*) = c, \qquad p_H^* = v_H(q_H^*), \qquad \text{and } \; \Pi^{\text{classic}} = \alpha \cdot (p_H^* - q_H^* \cdot c).$$

2.3 Preliminaries

Before proceeding to the main analysis, we present several preliminary results. First, observe that $\mathbf{R}(\cdot)$ affects the monopolist's optimization problem only through the ex ante expected reward $R \equiv \alpha \mathbf{R}(H) + (1-\alpha)\mathbf{R}(L)$. Thus, profits are identical for all reward schemes $\mathbf{R}(\cdot)$ that share the same expected value. Formally, this means:

Lemma 1 (Reward Reduction)**.** *If a menu of contracts* $((p_L, q_L), (p_H, q_H)) \in (\mathbb{R} \times \mathbb{R}_+)^2$ *and a reward scheme* $\mathbf{R}^{**} : \{L, H\} \to \mathbb{R}_+$ *solve* (1)*, then the same menu of contracts* $((p_L, q_L), (p_H, q_H))$ *and any reward scheme* $\mathbf{R} : \{L, H\} \to \mathbb{R}_+$ *with* $\mathbb{E}[\mathbf{R}] = \mathbb{E}[\mathbf{R}^{**}]$ *solve* (1)*.*

Despite being a simple observation, this result implies an important feature of the optimization problem faced by the firm. As long as the firm and the senders have the same expectation about the receivers' types, there is no reason for the firm to condition their payment on the purchased contract. Indeed, in Section 4.4 we show that if the senders have more accurate information about the receivers' types than the firm, the conclusion of Lemma 1 no longer holds. Thus, the detail of the optimal reward scheme crucially depends on the senders' knowledge. We relegate the analysis of this detail to Section 4.4, while here we consider senders who have the same information about the receiver's types as the firm does.

Plugging the sender's IC constraint into the objective function and noting that all senders share the same IC constraint, Lemma 1 allows us to simplify the problem as follows:

$$\Pi^* = \max_{p_L, p_H \in \mathbb{R},\, q_L, q_H \geq 0,\, R \in \mathbb{R}_+} \quad N \cdot \underbrace{G\left(r\left(\alpha + (1-\alpha) \cdot \mathbf{1}_{\{q_L > 0, v_L(q_L) \geq 0\}}\right) + R\right)}_{\text{probability of talking}} \cdot \\ \left[\alpha \cdot (p_H - q_H \cdot c) + (1-\alpha) \cdot (p_L - q_L \cdot c) - R\right] \tag{3}$$

subject to the constraints (2). We prove the existence of a solution to this problem for right-continuous functions G with finitely many jumps to accommodate both homogeneous costs as in the main part of the paper and heterogeneous costs G as considered in Section 4.1. The existence of a solution is not immediate as the objective function is not necessarily continuous, but right-continuity of G with only finitely many jumps suffices to establish existence.

Proposition 1 (Existence)**.** *The maximization problem (3) subject to* (2) *has a solution.*

We denote the (non-empty) *set of solutions* to this problem by

$$\mathcal{S} \subseteq (\mathbb{R} \times \mathbb{R}_+)^2 \times \mathbb{R}_+.$$

Moreover, for any menu of contracts $((p_L, q_L), (p_H, q_H))$ satisfying (2), we denote the expected profits obtained by a receiver conditional on being informed by

$$\pi((p_L, q_L), (p_H, q_H)) \quad = \quad \alpha(p_H - q_H \cdot c) + (1 - \alpha)(p_L - q_L \cdot c).$$

The monopolist can always choose not to sell to anyone and attain zero profits, i.e., $\Pi^* \geq 0$. Furthermore, whenever $\Pi^* = 0$ the seller can attain the maximum by inducing no sender to talk. This can be done by offering unacceptable contracts to receivers and no rewards.[12] We, thus, focus the characterization of optimal menu of contracts and rewards programs on the case when $\Pi^* > 0$.[13] The following lemma summarizes some basic properties of optimal menus of contracts.

Lemma 2. *If* $\Pi^* > 0$ *and* $((p_L, q_L), (p_H, q_H), R) \in \mathcal{S}$*, then:*

(i) **Low types don't pay:** $q_L \in \{0, \underline{q}\}$ *and* $p_L = 0$.[14]

(ii) **No distortions at the top:** $q_H = q_H^*$.

(iii) **No free contracts:** *If* $q_L = 0$*, then* $p_H = p_H^*$.

(iv) **Free contracts:** *If* $q_L = \underline{q}$*, then* $p_H = p_H^* - \underbrace{v_H(\underline{q})}_{\text{information rent}} \equiv \tilde{p}_H^*$.

Intuitively, the only benefit of selling to L-type receivers is that it increases the probability of the receiver using the product. Consequently, if a positive quantity is sold to L-type receivers, then it must be just enough to incentivize usage but no more. Moreover, the participation constraint of the L-type must be binding (as in Maskin and Riley (1984)).

[12] Note that if there is a positive mass of senders with $\xi = 0$, then by Assumption 3 the seller can attain strictly positive profits by only selling to H-receivers and offering no reward.

[13] In part 1 of Theorem 1, we give a necessary and sufficient condition for $\Pi^* > 0$ to hold.

[14] The proof in the Appendix shows that we do not need to restrict prices to be nonnegative in order to obtain this result.

Similarly, there are no distortions at the top. Parts (iii) and (iv) follow because the incentive compatibility constraint of H-type receivers must be binding.

Lemma 2 restricts the set of possible optimal contracts significantly. In particular, it uniquely pins down the price offered to low types and the quantity offered to high types whenever $\Pi^* > 0$. At a price of zero for low types, the seller either chooses $q_L = 0$ (*no free contracts*) or $q_L = \underline{q}$ (*free contracts*). A full characterization of optimal contracts requires us to characterize the optimal reward scheme R and whether free contracts are optimal for the monopolist. These choices depend on the parameters that have not been used so far: the cost structure, the magnitude of externalities, and the composition of different types of buyers.

3 Main Analysis

This section assumes that the cost of talking is homogeneous and equal to $\bar{\xi} > 0$ for all senders, i.e.,

$$G(\xi) = \mathbf{1}_{\{\bar{\xi} \leq \xi\}}. \tag{4}$$

This simple case allows us to illustrate the main trade-offs. As a robustness check, the Online Appendix deals with the case of heterogeneous costs in detail. We summarize the main insights of that analysis in Section 4.

3.1 Characterization of Optimal Scheme

We characterize the optimal contracts in steps. First, we characterize the optimal referral reward scheme given a menu of contracts satisfying (2) (Lemma 3). Then, we solve for the optimal menu of contracts (Lemma 4) and finally, use these optimal contracts to derive the optimal reward using Lemma 3 (Theorem 1).

With homogeneous costs of talking, if $r\left(\alpha + (1-\alpha)\cdot \mathbf{1}_{\{q_L>0, v_L(q_L)\geq 0\}}\right) + R \geq \bar{\xi}$, then for any menu of contracts satisfying the constraints (2), profits are given by $\pi((p_L, q_L), (p_H, q_H)) - R$. Otherwise, profits are zero. Thus, if incentivizing WoM is not more expensive than the expected profits, the monopolist would like to pay senders just enough to make them talk. The following lemma formalizes this intuition. Let

$$R^{**}((p_L, q_L), (p_H, q_H)) \quad = \quad \max\left\{\bar{\xi} - \underbrace{r \cdot \left[\alpha + (1-\alpha)\cdot \mathbf{1}_{\{q_L>0, q_L \geq \underline{q}\}}\right]}_{\text{expected externalities}}, 0\right\}. \tag{5}$$

Lemma 3 (Referral Program)**.** *Suppose G is given by* (4)*. Given contracts (p_L, q_L) and (p_H, q_H) satisfying* (2) *and $v_H(q_H) \geq 0$, the optimal referral reward is unique as long as $R^{**}((p_L, q_L), (p_H, q_H)) < \pi((p_L, q_L), (p_H, q_H))$ and is given by $R^{**}((p_L, q_L), (p_H, q_H))$.*

Using Lemma 2 and the formula of the optimal reward function R^{**} in Lemma 3, we can determine whether it is optimal to offer free contracts or not, which then pins down the full optimal menu of contracts.

In interpreting the full characterization, it is instructive to understand what the cost of offering free contracts is. It is given by the information rent that the firm needs to pay to v_H-buyers (pertaining to the share α of the receivers) and by the cost of producing the free product (pertaining to the share $1-\alpha$ of the receivers). The following variable quantifies the overall cost of free contracts:

$$CF^* \quad \equiv \quad \alpha \quad \cdot \underbrace{v_H(\underline{q})}_{\text{information rent}} \quad + (1-\alpha)\cdot \quad \underbrace{c\cdot \underline{q}}_{\text{production cost of free product}} \quad . \tag{6}$$

Using this variable, let us first provide a heuristic argument: In order for free contracts to be optimal, this cost has to be outweighed by the benefit generated by providing the product to low types, i.e.,

$$CF^* \leq (1-\alpha) r, \tag{7}$$

or equivalently $\frac{CF^*}{1-\alpha} \leq r$. Notice that $\frac{CF^*}{1-\alpha}$ represents the "break-even externalities" necessary to compensate for the cost of free contracts. Moreover, $\frac{CF^*}{1-\alpha}$ is increasing in α. The average profit generated by a receiver if free contracts are offered can be written as

$$\pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*)) = \Pi^{\text{classic}} - CF^*$$

The following result shows that, with additional boundary conditions, (7) is also sufficient to guarantee optimality of free contracts. We denote the set of optimal q_L by Q_L^{**}.

Lemma 4 (Free Contract)**.** *Suppose G is given by* (4)*. Whenever $\Pi^* > 0$, an optimal contract to the type-L receiver must satisfy the following:*

(i) Let $r \in [\frac{\bar{\xi}}{\alpha}, \infty)$. Then, $Q_L^{**} = \{0\}$ (i.e., it is not optimal to provide free contracts).

(ii) Let $r \in [\bar{\xi}, \frac{\bar{\xi}}{\alpha})$.

1. **(Free contracts)** $\underline{q} \in Q_L^{**}$ *if and only if*

$$\underbrace{\bar{\xi} - \alpha r}_{\text{reward w/o free contract}} \geq CF^*. \tag{8}$$

2. **(No free contracts)** $0 \in Q_L^{**}$ *if and only if* $\bar{\xi} - \alpha r \leq CF^*$.

(iii) Let $r \in [0, \bar{\xi})$.

1. **(Free contracts)** $\underline{q} \in Q_L^{**}$ *if and only if* $r \geq \frac{CF^*}{1-\alpha}$.

2. **(No free contracts)** $0 \in Q_L^{**}$ *if and only if* $r \leq \frac{CF^*}{1-\alpha}$.

The intuition for this lemma is the following. First, there is no need for the seller to provide any incentives for WoM (i.e., $q_L = 0$) if the cost of talking $\bar{\xi}$ is smaller than the lowest expected externalities αr because in that case people talk anyway (Lemma 4 (i)). If $r \in [\bar{\xi}, \frac{\bar{\xi}}{\alpha})$ (Lemma 4 (ii)), then the cost of talking is larger than αr, but free contracts can boost the expected externalities to $r \geq \bar{\xi}$. Then, free contracts are used whenever the referral reward that the seller had to pay without free contracts $\overline{\xi} - \alpha r$ is larger than the cost of offering a free contract CF^* which is the sum of the information rent and cost of producing $\underline{q}$. Note that in this case, whenever free contracts are offered, the optimal reward is zero by Lemma 3. Finally, for high costs of talking $\overline{\xi} > r$ (Lemma 4 (iii)), by Lemma 3 the seller pays a reward as long as the optimal reward does not exceed expected profits. If free contracts are offered, the expected externalities can be increased by $(1-\alpha)r$. Hence, free contracts are offered only if this benefit exceeds the cost of production and the information rent so that $r \geq \frac{CF^*}{1-\alpha}$ as explained above.

Lemmas 2, 3 and 4 pave the way for a full characterization of the optimal menu of contracts and reward scheme summarized in the following theorem. It shows that the optimal incentive scheme depends on the market structure given by parameters such as the cost of production c, the externalities r, the cost of talking $\overline{\xi}$, and the fraction of H-type receivers α.

Theorem 1 (Full Characterization)**.** *Suppose G is given by* (4)*.*

1. **(Positive profits)** $\Pi^* > 0$ *if and only if*

$$\overline{\xi} < \max\left\{\Pi^{classic} - CF^* + \min\{r, \overline{\xi}\},\ \Pi^{classic} + \alpha r\right\}. \tag{9}$$

For the following cases, assume that (9) *is satisfied:*

2. **(Free vs. no free contracts)** *There exists* $((0,\underline{q}),(\tilde{p}_H^*, q_H^*), R) \in \mathcal{S}$ *for some* R *if and only if* $r \in \left[\frac{CF^*}{1-\alpha}, \frac{\xi - CF^*}{\alpha}\right]$.[15]

3. **(Rewards vs. no rewards)**

   (a) **(With free contracts)** *If* $r \in [\frac{CF^*}{1-\alpha}, \frac{\xi - CF^*}{\alpha}]$*, then* $((0,\underline{q}),(\tilde{p}_H^*, q_H^*), R) \in \mathcal{S}$ *with* $R > 0$ *if and only if* $r < \bar{\xi}$*, and*

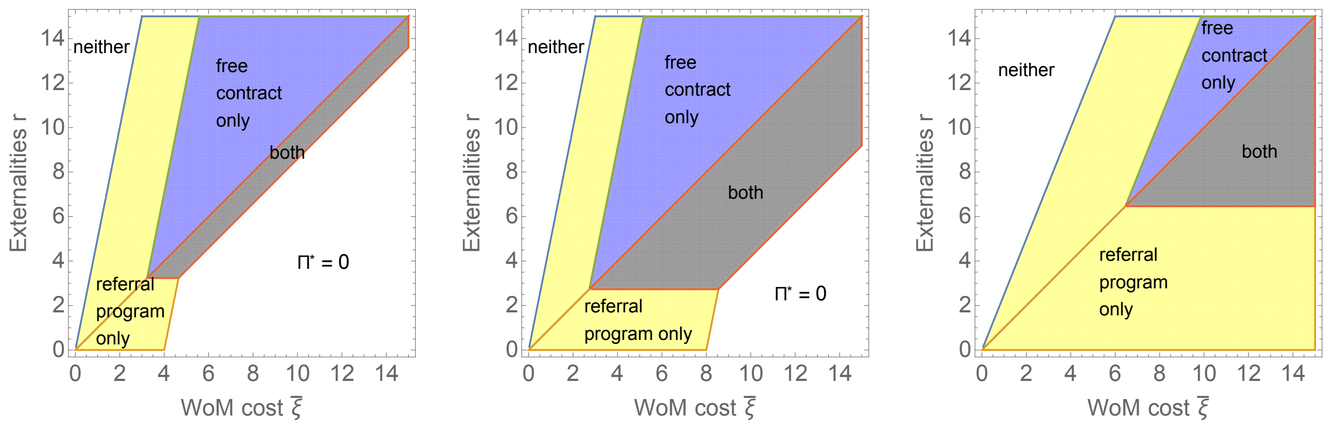
[15] If $\frac{CF^*}{1-\alpha} > \frac{\xi - CF^*}{\alpha}$, then $[\frac{CF^*}{1-\alpha}, \frac{\xi - CF^*}{\alpha}] = \emptyset$.

(b) **(With no free contracts)** *If* $r \notin [\frac{CF^*}{1-\alpha}, \frac{\xi - CF^*}{\alpha}]$*, then* $((0,0),(p_H^*, q_H^*), R) \in \mathcal{S}$ *with* $R > 0$ *if and only if* $r < \frac{\bar{\xi}}{\alpha}$.

First, it is straightforward that the monopolist should provide no incentives for WoM either if senders talk anyway because the cost of talking is small (i.e., $\overline{\xi} < \alpha r$) or if it is too expensive because the cost of talking $\overline{\xi}$ is too large relative to its benefits given in (9). A necessary condition for free contracts to be optimal is that r is large enough (i.e., $r > \frac{CF^*}{1-\alpha}$). An immediate implication is that without any externalities, free contracts are of no value to the seller. At the same time, free contracts are more effective to encourage WoM than rewards only if the cost of talking $\overline{\xi}$ is sufficiently large relative to r (i.e., $\overline{\xi} > CF^* + \alpha r$ which is derived from the upper bound of r in part 2 of Theorem 1). Otherwise, it is cheaper to pay a small reward for talking. We discuss comparative statics with respect to α and r in the next section.

Figure 1 illustrates the different regions in the $(\overline{\xi}, r)$-space characterized in Theorem 1 for $v_H(q) = 2\sqrt{q}$, $\underline{q} = 20$ (i.e., $v_H(\underline{q}) \simeq 8.94$), and for different production costs c and fraction of H-type receivers α. The left panel shows the different regions for $\alpha = 0.2$ and $c = 0.05$ (i.e., $q_H^* = 400$, $p_H^* = 40$), while the middle panel assumes lower cost of production $c = 0.025$ (i.e., $q_H^* = 1600$, $p_H^* = 80$). Comparing these two figures, one can see how low marginal cost of production c gives the seller incentives to encourage WoM (with free contracts and/or rewards) for high costs of talking ξ.

The rightmost panel of Figure 1 shows the different regions for a larger fraction of H-type receivers ($\alpha = 0.4$). We can think of markets with such high α as *mass markets*, in contrast to *niche markets* with small fractions α of H-type buyers. The comparison of the two right panels indicates that in mass markets free contracts are not optimal for relatively small externalities r and cost of talking $\overline{\xi}$.



(a) Niche market ($\alpha = 0.2$) with $c = 0.05$

(b) Niche market ($\alpha = 0.2$) with $c = 0.025$

(c) Mass market ($\alpha = 0.4$) with $c = 0.025$

Figure 1: Equilibrium Regions in the $(\bar{\xi}, r)$-space

3.2 Comparative Statics and Discussion

Motivated by the last observation about mass versus niche markets, here we fix $\bar{\xi}$ and analyze the different implications for the menu of optimal contracts and reward scheme as the market size α varies. Our model predicts a pricing pattern consistent with those that we observe in the real world.

Proposition 2 (Market Structure and Free Contracts)**.** *Suppose G is given by* (4)*.*

(i) Consider two markets that are identical to each other except for the share of H-types, denoted α_1 and α_2. Suppose that free contracts are offered under an optimal scheme in the market with α_1, $\Pi^* > 0$ in the market with α_2, and $\alpha_2 < \alpha_1$. Then, free contracts are offered under any optimal scheme in the market with α_2.

(ii) Suppose $v_H(\underline{q}) + r > c\underline{q}$. Then, $\alpha > \frac{r - c\underline{q}}{v_H(\underline{q}) + r - c\underline{q}}$ $(\Leftrightarrow r < \frac{CF^*}{1-\alpha})$ implies that free contracts are never offered under any optimal scheme.

This proposition shows that the monopolist should encourage WoM in a market with a small fraction α of H-type buyers as long as the market is profitable enough, i.e., $\Pi^{**} > 0$. Intuitively, if there are many H-types, the seller is better off paying a reward because free contracts do not increase the probability of purchase by much. The exact trade-off is determined by the comparison of the information rent and the per-low-type surplus $r - c\underline{q}$

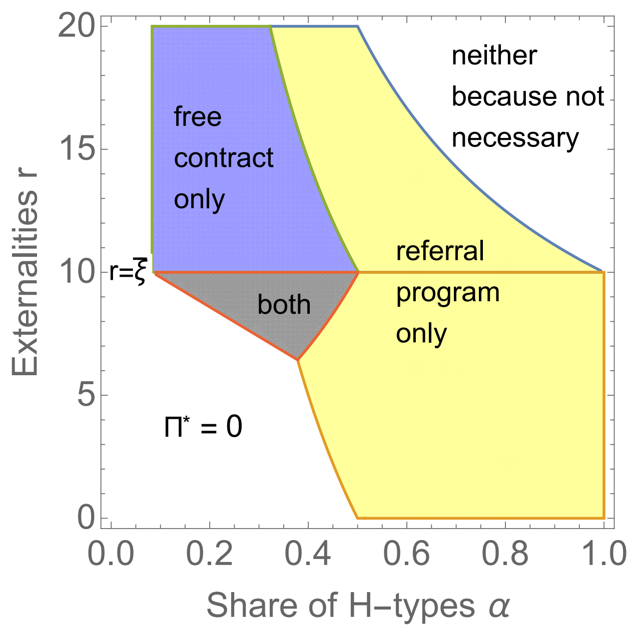


Figure 2: Equilibrium Regions in the (α, r)-space

that the seller can extract. The cutoff for α is increasing in this rent while decreasing in the information rent.

Figure 2 illustrates the different regions in the (α, r)-space given the same parameters as in Figure 1. It shows that free contracts are only optimal for small fractions α of H-buyers. However, if there are too few H-buyers (i.e., $\alpha < 0.08\ldots$), then profits generated become too small to make it worthwhile to encourage WoM (i.e., $\Pi^* = 0$).[16] With small externalities r, senders have little innate benefit from WoM, so the lower bound of α above which the profit is positive is large.

These findings are consistent with the observation that digital service providers with small production costs who successfully offer free contracts (e.g., Dropbox or Skype), have a large number of free users. Moreover, free contracts are combined with a reward program, if the externalities are not large (as in Dropbox: one may use it for oneself to store files and access them from multiple computers, or share files with others), while only free contracts are offered if the externalities are large (as in Skype: any usage generates externalities). In

[16]This region disappears with heterogeneous priors as we show in the Online Appendix.

Externalities	$r < \frac{CF^*}{1-\alpha}$	$\frac{CF^*}{1-\alpha} < r < \bar{\xi}$	$\bar{\xi} < r < \frac{\bar{\xi}-CF^*}{\alpha}$	$\frac{\bar{\xi}-CF^*}{\alpha} < r < \frac{\bar{\xi}}{\alpha}$	$\frac{\bar{\xi}}{\alpha} < r$
Referral rewards	Yes	Yes	No	Yes	No
Free contracts	No	Yes	Yes	No	No
Profit	Positive or zero	Positive or zero	Positive	Positive	Positive

Table 1: Comparative Statics with respect to r when $\bar{\xi} < \frac{CF}{1-\alpha}$. The use of referral rewards and free contracts is conditional on the firm generating positive profits.

contrast, transportation services such as Amtrak or Uber that solely rely on referral rewards programs would correspond to monopolists facing high α and low r, as many customers would be willing to pay for such services and those services would not be subject to significant externalities.[17]

One might think that the smaller the externalities are, the more likely rewards are used. Figure 2 illustrates that this type of comparative statics fails for externalities. For example, at $\alpha = 0.4$, referrals are used when $r = 20$ but not when $r = 12$. The reason is that (i) when r is high, only one of free contracts and referrals suffices to incentivize the senders, i.e., these two are substitutes, and (ii) the cost of offering free products CF^* is constant across r's while the optimal reward monotonically decreases with r. Thus, conditional on offering free contracts being sufficient to encourage WoM (i.e., $r \geq \overline{\xi}$), offering free contracts is more cost-saving for smaller r while rewards are more cost-saving for larger r. Table 1 summarizes the different regions as functions of r for the case in which $\bar{\xi} < \frac{CF}{1-\alpha}$.[18]

In the following proposition, we make the claim in (i) clearer by defining what we mean by the two strategies being "substitutes."

Proposition 3 (Substitutes)**.** *Referrals and free contracts are strategic substitutes as long*

[17]Note that the fraction of the consumers purchasing free contracts is an endogenous variable, and one might think that our association of observable fractions for these real products to the exogenous parameter α is not justifiable. However, such association is justified because the map from consumer types to the choices of contracts is one-to-one given that free contracts are used. That is, if a positive fraction of consumers purchases free contracts, then within our model, such a fraction is exactly equal to $1 - \alpha$. Yet, it may be hard to empirically test our predictions for firms that do not offer free contracts given that absent free contracts we do not observe α.

[18]If this condition is not satisfied, some regions cease existing.

as it is optimal to have a referral program without free contracts, i.e.,

$$R^{**}((0,0),(p_H^1,q_H^1)) > R^{**}((\underline{q},0),(p_H^2,q_H^2)) \tag{10}$$

for all $(p_H^1,q_H^1),(p_H^2,q_H^2) \in \mathbb{R}_0 \times \mathbb{R}$ *such that (i)* $R^{**}((0,0),(p_H^1,q_H^1)) < \pi((0,0),(p_H^1,q_H^1))$ *and (ii) both menu of contracts* $((0,0),(p_H^1,q_H^1))$, $((0,\underline{q}),(p_H^2,q_H^2))$ *satisfy* (2).

Intuitively, a sender is willing to talk only if the expected externalities from talking are large enough. Thus, the monopolist can either directly pay the sender or increase the likelihood of successful referrals by offering free contracts to L-type receivers. Put differently, free contracts (paying the receiver) can be a substitute for reward payments (paying the sender). Note that there are situations where it is too expensive to incentivize WoM with rewards programs only (such that $R^{**}((0,0),(p_H,q_H)) = 0$), but the seller might benefit from a positive reward R in combination with free contracts. In that case, (10) is not satisfied.

In order to see the implication of the substitution result on the optimal contract and reward scheme, Figure 3 depicts the reward under the optimal menu of contracts as a function of parameters α and r. In Figure 3-(a), there is a discontinuous upward jump at around $\alpha = 0.4$. That is, at the point where the parameter region changes from the one where both free contracts and referral rewards are used to the one where only a referral program is used, the amount of the optimal reward goes up. This is precisely because of the substitution effect: Because the free contracts are dropped, the reward has to increase. Note that the same pattern appears in Figure 3-(b) that depicts the optimal reward as a function of the externalities r. In that graph, there is a discontinuous downward jump at around $r = 8$ where the parameter region changes from the one where only a referral program is used to the one where both free contracts and referral rewards are used.

Note that the optimal amount of reward goes down as α goes up or r goes up in the region where only a referral program is used. This is because high α and high r means a higher expected benefit from talking with everything else equal, so there is less need to provide

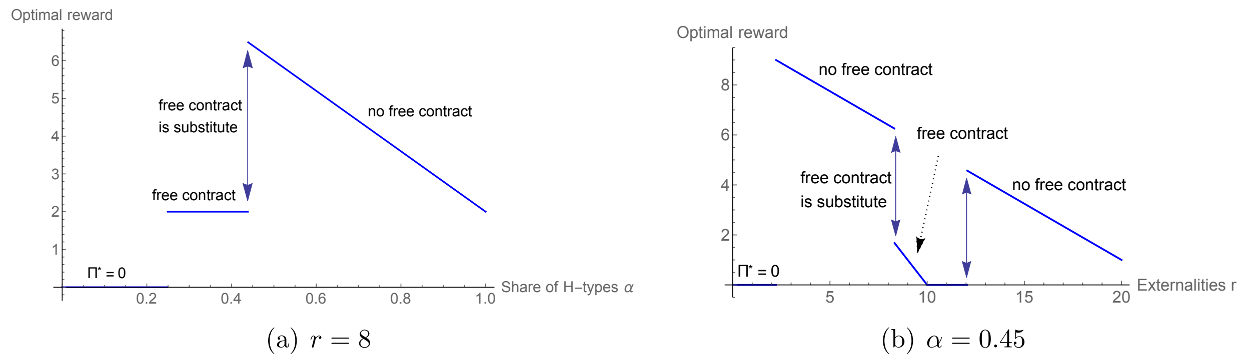


Figure 3: Rewards under the Optimal Scheme

a large reward. On the other hand, the optimal reward is constant in α but decreasing in r in the region where both free contracts and referral rewards are used. It is constant in α because the receivers will be using the product (once informed) under provision of free contracts, so the expected benefit from talking does not depend on α. It is decreasing in r for the same reason as for the region where only a referral program is used.

4 Discussion

In this section we discuss various extensions and their implications, as well as the social planner's problem.

4.1 Heterogeneous WoM Cost

In Section 3, we have entirely focused on homogeneous costs of talking (i.e., G follows (4)), in order to emphasize the core trade-off faced by a firm when encouraging senders to engage in WoM. In the Online Appendix, we consider an extension in which different senders have different costs of talking. With heterogeneous costs of talking, the optimal reward scheme is more complicated as it can be used to *fine-tune the amount of WoM*, while with homogeneous costs either everyone or no one talks. We analyze the optimal scheme for a fairly general class of cost distribution G, and discuss how our results from Section 3 change. Here, we summarize the main findings of that section.

We show that the results from Section 3 are robust in the following sense. Free contracts are not optimal for large α because in that case the benefit of free contracts given by $(1-\alpha)r$ is small compared to the cost CF^*. Referrals and free contracts remain strategic substitutes. We also show how the homogeneous cost case can be thought of as the limit of models with heterogeneous costs.

New insights can be derived in the heterogeneous cost model with respect to the reward scheme. The optimal reward scheme is not constant in α when a free contact is offered (as it is when G follows (4)), but is increasing in α. The reason is that expected profits are higher with higher α and hence, the seller has a stronger incentive to increase WoM. If no free contracts are offered, in addition to the aforementioned effect, there is an opposing effect (that is present also with homogeneous costs), as the seller only needs to pay less to senders if the expected externalities are large in order to induce the same number of senders to talk. Thus, if no free contracts are offered the effect of α on rewards is ambiguous, where rewards are decreasing in α if costs are sufficiently homogeneous.

4.2 Two-Sided Externalities

In the main analysis we assumed that only the senders receive externalities, and claimed that even if we assumed the receivers would receive externalities as well, the essence of the analysis would not change. The goal of this subsection is to make this formal. Consider a model as in Section 2, with an additional feature that if receiver i uses the product, she receives externalities r. In this model, for each $\theta \in \{H, L\}$, if a type-θ receiver uses quantity q, she experiences utility $v_\theta(q) + r$.

Note that this is a change that shifts the valuation functions by a constant, i.e., they change from $v_\theta(q)$ to $v_\theta(q) + r$ for each $\theta = H, L$. Hence, it does not alter the nature of the optimal contract scheme *under each fixed* r, assuming that our restrictions are met for the new valuation functions. This implies that all comparative statics with respect to parameters that are not r (e.g., Proposition 2) are robust. Below we show that our main comparative statics with respect to r (provided in Theorem 1) goes through as well.[19]

Note that Theorem 1 states that the use of free contracts is optimal if and only if the condition $r \in \left[\frac{CF^*}{1-\alpha}, \frac{\xi - CF^*}{\alpha}\right]$ is met. Then, the use of rewards is determined by conditions given by the bounds independent of the size of r (the conditions are $r < \bar{\xi}$ in the presence of free contracts and $r < \frac{\bar{\xi}}{\alpha}$ otherwise, and $\bar{\xi}$ and $\frac{\bar{\xi}}{\alpha}$ do not depend on r). It is immediate that the same characterization goes through in our modified model, but now the size of CF^* depends on r. If we show that CF^* is nonincreasing and $CF^* + \alpha r$ is nondecreasing in r, then the region of r such that free contracts are used is still given by a convex interval, guaranteeing that the essence of the comparative statics does not change. We first show that CF^* is strictly decreasing in r. To show this, let us write down the modified CF^* as follows:

$$CF^*(r) = \alpha(v_H(\underline{q}(r)) + r) + (1-\alpha)c\underline{q}(r),$$

where $CF^*(r)$ and $\underline{q}(r)$ denote the cost of free contracts under r and the break-even quantity for low-types under r (i.e., $v_L(\underline{q}(r)) + r = 0$), respectively. It is immediate that the second term is strictly decreasing in r because $v_L'(q)$ is strictly increasing in q and thus $\underline{q}(r)$ is strictly decreasing in r. The first term is strictly decreasing in r for the following reason: Take r and r' with $r < r'$. Then, by the assumption that $v_H'(q) > v_L'(q)$ and the definition of the $\underline{q}(\cdot)$ function, it must be the case that:

$$(v_H(\underline{q}(r'))+r')-(v_H(\underline{q}(r))+r) < \big(v_L(\underline{q}(r')) - v_L(\underline{q}(r))\big)+(r'-r) = ((-r') - (-r))+(r'-r) = 0$$

Overall, $CF^*(r)$ is strictly decreasing in r. We next show that $CF^*(r) + \alpha r$ is strictly increasing in r under an additional assumption about the valuation functions. Specifically, suppose that $2v_L'(q) > v_H'(q) + \frac{1-\alpha}{\alpha}c$ for all $q > 0$. That is, the marginal values of the two types are not too different form each other, which ensures that the information rent $v_H(\underline{q}(r))$ does not vary too much with r. Then, taking the first-order condition of CF^* with respect

[19] We keep assuming that our restrictions are satisfied after the shifts of the valuation functions.

to r and by noting $\underline{q}'(r) = -\frac{1}{v_L'(\underline{q}(r))}$ (by the Implicit Function Theorem), one can show that $CF^*(r) + \alpha r$ is strictly increasing in r. All in all, free contracts are used if and only if r is in a convex interval.

Note that this analysis provides an interesting observation that the cost of free contracts decreases in the size of externalities because both the production cost and the information rent decrease. The reason is that if low types receive externalities it becomes easier for the firm to make them willing to use the product (implying low production cost) and high types have less incentives to switch to the low-type contract at such a level of quantity provided to low types (implying lower information rent).

To sum up, the model of two-sided externalities provides qualitatively equivalent comparative statics as our main model with one-sided externalities.

4.3 Quantity-Dependent Externalities

The main analysis is based on a model in which the magnitude of externalities is captured by a single parameter r. As Theorem 1 shows, this is the key parameter that determines the optimal scheme. However, one can imagine that a Dropbox user who wants to refer his co-author receives higher positive externalities from joint usage if the co-author uses Dropbox more. The objective of this section is to formalize the idea of quantity-dependent externalities and discuss how such dependencies affect our predictions.

To this end, consider a function $\bar{r} : \mathbb{R}_+ \to \mathbb{R}_+$ that assigns to each quantity level consumed the value of externalities generated. We employ the normalization that $\bar{r}(0) = 0$. Note that our main model corresponds to the case in which $\bar{r}(q) = r$ for all $q > 0$. In this section we assume that $\bar{r}$ is differentiable, strictly concave, $\bar{r}'(q) > 0$ for all $q \geq 0$ and $\lim_{q\to\infty} \bar{r}'(q) = 0$.

Fix an optimal scheme $((\bar{p}_L^*, \bar{q}_L^*), (\bar{p}_H^*, \bar{q}_H^*), \overline{R}^*)$. Then, the L-type's PC constraint and the H-type's IC constraint must be binding. First, consider the case when the sender's IC constraint is binding. In that case, (generically) positive rewards are being paid. Then, if a contract is offered to the low types ($\bar{q}_L^* > 0$), then the optimal scheme must satisfy the following first-order condition:

$$\alpha(v'_H(\bar{q}^*_H) - c + \bar{r}'(\bar{q}^*_H)) \;\; = \;\; 0$$

and $\bar{q}^*_L \in \{0, \underline{q}\}$ (as in the main model) if

$$(1-\alpha)(v'_L(q_L) - c + \bar{r}'(q_L)) + \alpha(v'_L(q_L) - v'_H(q_L)) \;\; < \;\; 0 \tag{11}$$

holds for $q_L = \underline{q}$, and $\bar{q}^*_L$ satisfies the above inequality with equality otherwise.[20] For simplicity, we focus the discussion on the case when the inequality in (11) is satisfied for $q_L = \underline{q}$.

Otherwise, the contract has a positive price. If low types are not served under the optimal contract scheme, then only the first first-order condition need be satisfied. Thus, as in the main model, there are only three possible levels of realized externalities corresponding to the three contracts that the firm optimally chooses conditional on rewards being paid, $\bar{r}(\bar{q}^*_H) =: r_H$, $\bar{r}(\underline{q}) =: r_L$ and $\bar{r}(0) = 0$. Note that in this case, $q^*_H \leq \bar{q}^*_H$ holds because $\bar{r}'(\bar{q}^*_H) > 0$ and v'_H is decreasing.

If the sender's IC constraint is not binding, then the sender's IC can be ignored and thus, the optimal contract is the same as in the main model, and in particular, $\bar{q}^*_H = q^*_H$. Let us denote the externalities received if the high type's contract is purchased by $r_h := \bar{r}(q^*_H)$.

Here we consider how the conditions for offering free contracts change. In the absence of free contracts, expected externalities are given by αr_H, while in the presence of free contracts, expected externalities are given by $\alpha r_H + (1-\alpha) r_L$. Now, consider part 2 of Theorem 1. It says that, for free contracts to be used in the optimal scheme, two conditions have to be met: $r(1-\alpha) \geq CF^*$ and $\xi - \alpha r \geq CF^*$. The first inequality says that the cost of free contracts has to be no more than the increment of the expected externalities. The second says that it has to be no more than the rewards necessary to be paid to compensate for the difference between the cost of talking and the externalities that are generated anyway by high types,

[20] The solution exists and is unique as we assume $\bar{r}$ is strictly concave and the limit of its slope is zero.

in the absence of free contracts. Since the first inequality automatically holds when the sender's IC constraint does not bind, and the second inequality automatically holds when the sender's IC constraint binds, these conditions can be rewritten as:

$$r_L(1-\alpha) \geq CF^* \text{ and } \xi - \alpha r_h \geq CF^*.$$

Since CF^* is unchanged, these conditions imply that low externalities for low types and high externalities for high types both reduce the set of parameters for which free contracts are optimally offered. Thus, free contracts can be optimal only if the dependence of the magnitude of externalities does not vary too much with the quantity consumed by the receivers. Our main analysis corresponds to the (extreme) case with constant $\bar{r}$ functions, and hence best captures the role of free contracts.

4.4 Informed Senders

To simplify the analysis, in the main analysis we assume that each sender has the same information about the type of his receiver as the firm. However, in some markets one can imagine that senders have better information about their friends' willingness to pay than the firm. The objective of this section is to consider a model that accommodates this possibility, and to discuss robustness of and difference from the results of the main analysis. Specifically, let us assume that each sender independently observes a signal $s \in \{s_L, s_H\}$ about his receiver. If the receiver's type is $\theta = H$, the sender sees a signal $s = s_H$ with probability $\beta \in \left(\frac{1}{2}, 1\right)$, and if the receiver's type is $\theta = L$, the sender sees a signal $s = s_H$ with probability $1-\beta$.[21] Thus, by Bayes rule, a sender who has received a signal s_H believes that the probability of facing a H-type receiver is $\alpha_H = \frac{\alpha\beta}{\alpha\beta+(1-\alpha)(1-\beta)}(> \alpha)$, while a sender who has received a signal s_L, instead believes that the probability of facing a H-type receiver is $\alpha_L = \frac{\alpha(1-\beta)}{\alpha(1-\beta)+(1-\alpha)\beta}(< \alpha)$.

[21] If $\beta = \frac{1}{2}$ was the case, then senders and the firm would have exactly the same information about receivers. Our main model corresponds to this case.

How does the firm's optimization problem change? The firm's objective function is a weighted sum of the profit generated by WoM of senders who have received a high signal and the profit generated by WoM of senders who have received a low signal. The two profit functions are as in (1) with the fraction of high valuation receivers being α_H and α_L, respectively. More precisely, a fraction $\alpha\beta + (1-\alpha)(1-\beta)$ of senders have received a high signal s_H and the expected profits generated by those senders is just (1) with the fraction of H-type receivers being α_H. A fraction $\alpha(1-\beta)+(1-\alpha)\beta$ of senders has received a low signal and the profit generated by those senders is (1) with the fraction of H-type receivers being α_L. Note that the receivers' constraints remain unchanged. However, the firm now faces two IC constraints for the senders - one for the senders who observed s_H and one for the senders who observed s_L.

An important difference to the model we consider in the main part is that Lemma 1 is not valid anymore as the firm can utilize the informational differences with the reward scheme.

Proposition 4 (Rewards with informed senders)**.** *1. Suppose that all senders choose "Refer" under the optimal scheme.*

(a) If the firm does not offer free contracts, then the optimal reward scheme $\mathbf{R}$ *satisfies* $\mathbf{R}(H) \leq \mathbf{R}(L)$ *with the inequality being strict if* $r \in (0, \frac{\overline{\xi}}{\alpha_L})$.[22]

(b) If the firm offers free contracts, then the optimal reward scheme $\mathbf{R}$ *satisfies* $\mathbf{R}(H) = \mathbf{R}(L) = \max\{\overline{\xi} - r, 0\}$.

2. Suppose that senders who received s_H *choose "Refer" but other senders choose "Not" under the optimal scheme.*

(a) If the firm does not offer free contracts, then there exists an optimal reward scheme $\mathbf{R}$ *such that* $\mathbf{R}(H) > \mathbf{R}(L) = 0$*. Moreover, any optimal reward scheme* $\mathbf{R}$ *satisfies* $\mathbf{R}(H) > \mathbf{R}(L) - r$.

[22] $\mathbf{R}(H) = \overline{\xi} - r < \mathbf{R}(L) = \overline{\xi}$ for $\overline{\xi} \geq r$ and $\mathbf{R}(H) = 0 \leq \mathbf{R}(L) = \max\left\{\frac{\overline{\xi} - \alpha_L r}{1-\alpha_L}, 0\right\}$ for $\overline{\xi} < r$.

(b) *If the firm offers free contracts, then there exists an optimal reward scheme* $\mathbf{R}$ *such that* $\mathbf{R}(H) > \mathbf{R}(L) = 0$*. Moreover, any optimal reward scheme* $\mathbf{R}$ *satisfies* $\mathbf{R}(H) > \mathbf{R}(L)$*.*

Each of the four cases arises given a nonempty parameter region that we compute in the proof of Proposition 5 in the Appendix. An important implication of this proposition is that, if the firm wants to incentivize all senders to talk, then she must pay *more* for referrals of L-type receivers than for H-type receivers because L-type senders' expected externalities are low. In contrast, if the firm is better off excluding senders who received signal s_L, then one optimal scheme only rewards referrals of premium users. Note that if the firm wants to induce s_L-senders to talk, it should also induce s_H-senders to talk because it is cheaper to provide incentives to s_H-senders and they talk to a better pool of receivers.

Solving the full problem is a daunting task because there are multiple cases to analyze depending on which type of senders are encouraged to talk. If the monopolist decides to encourage every sender to talk, the choice between free contracts and referral rewards can be tricky: offering free contracts can be very attractive in a market with fraction α_L of high types but not attractive in a market with fraction α_H of high types. As the firm cannot differentiate between buyers who have generated a high signal versus a low signal, it needs to trade off the benefits in both markets when deciding whether to offer free contracts. One can, however, easily derive the following results for the extreme cases:

Proposition 5 (Signal strength)**.** 1. *If* $\bar{\xi} - r < \alpha(p_H^* - cq_H^*)$*, then there exists* $\bar{\beta} < 1$ *such that for all* $\beta > \bar{\beta}$*, the unique optimal menu of contracts is given by* $((0,0),(p_H^*, q_H^*))$*, and there exists an optimal reward scheme* $\mathbf{R}$*, which satisfies* $\mathbf{R}(L) = 0$*. If* $\bar{\xi} - r \geq \alpha(p_H^* - cq_H^*)$*, then for any* $\beta \in (\frac{1}{2}, 1)$*, the firm cannot make positive profits.*

2. *Suppose that there exists a unique optimal menu of contracts* $((p_L, q_L), (p_H, q_H))$ *in the model without signals. Then, for all* $r \notin \left\{\frac{\bar{\xi}}{\alpha}, \frac{CF^*}{1-\alpha}, \frac{\bar{\xi}-CF^*}{\alpha}\right\}$*, there exists* $\bar{\beta} > \frac{1}{2}$ *such that for all* $\beta \in (\frac{1}{2}, \bar{\beta})$*, there exists a unique optimal menu of contracts and it is*

$((p_L, q_L), (p_H, q_H))$.

Part 1 shows that, if the signal strength β is too large, free contracts are not used by the seller. Part 2 then shows that the model we analyze in the main section without signals is reasonable when we think of the introduction of a new product category because in such a case β would be close to $\frac{1}{2}$.

4.5 Effect of Advertising

In this section, we investigate how the optimal incentive scheme changes if the firm can also engage in classic advertising. Formally, consider the situation in which the firm has an option to conduct costly advertising before WoM takes place. The firm spends $a \in \mathbb{R}_+$ for advertising and this is observed by all senders but not by any receivers. Then, each receiver independently becomes aware of the product prior to the communication stage with probability $p(a)$, where $p(0) = 0$ and $p(a) > 0$ for $a > 0$. The firm simultaneously chooses a menu of contracts, a reward scheme, and advertising spending. We assume that the sender does not observe whether the receiver is already aware of the product and only enjoys externalities if the receiver starts using the product *and* she engages in WoM (independently of whether the receiver learns through advertising and/or WoM) since otherwise she cannot know whether the receiver uses the product or not. The reward scheme is now a function $\mathbf{R} : \{L, H\} \times \{A, N\} \to \mathbb{R}_+$. Here, $\mathbf{R}(\theta, A)$ denotes the reward paid to the sender whose receiver purchases the contract offered to θ-types and becomes aware of the product through advertising. Similarly, $\mathbf{R}(\theta, N)$ denotes the reward paid to the sender whose receiver purchases the contract offered to θ-types and does not become aware of the product through advertising.[23]

Having completely specified the model with advertising, let us now analyze it. Note first that Lemma 2 again holds without any modification. Suppose now that the reward scheme $\mathbf{R}$

[23]We assume that the externalities r do not depend on a. Such dependence may arise if WoM is conducted with self-enhancement motive as in Campbell et al. (2015). In such a model, r would be decreasing in a, and advertising becomes an even less attractive option for the firm than in the current model.

and the advertising level a is part of the optimal scheme, and all senders choose Refer under such an optimal scheme. We assume $a > 0$ and derive a contradiction. To show this, consider the following modification of the scheme. First, let $R \equiv \alpha\,(p(a)\mathbf{R}(H,A) + (1-p(a))\mathbf{R}(H,N)) + (1-\alpha)\,(p(a)\mathbf{R}(L,A) + (1-p(a))\mathbf{R}(L,N))$ be the expected reward, and construct a new reward scheme $\mathbf{R}'$ such that $\mathbf{R}'(\theta, x) = R$ for all $\theta = H, L$ and $x = A, N$. As in Lemma 1, this new scheme also satisfies the constraints and gives rise to the same expected profit, so it is optimal too. Now, consider changing $a > 0$ to a new advertising level $a' = 0$. With the new scheme $(\mathbf{R}', a')$, the constraints are still satisfied; in particular all the senders choose Refer. Also, the expected profit to the monopolist increases by $a > 0$. This contradicts the assumption that the original scheme with $(\mathbf{R}, a)$ is optimal. All in all, this argument implies that either (i) the firm chooses a positive advertising level and no WoM takes place or (ii) WoM takes place and $a = 0$. Note that, in case (i), compared to the model in Section 2, advertising either substitutes WoM or allows the firm to inform some receivers if encouragement of WoM was too expensive.

4.6 Multiple Senders per Receiver

In the main model, we consider a stylized network structure between senders and receivers, i.e., receiver i is connected only to sender i, and *vice versa*. In reality, however, it is possible that a receiver is connected to multiple potential senders of the same information. Similarly to the discussion in Section 4.5 where the receiver can learn from an advertisement, a receiver has multiple sources of information if there are multiple senders. Such a situation can arise when senders and receivers are connected through a general network structure.

In this section we discuss how the predictions change when there are multiple senders per receiver. To make our point as clear as possible, let us assume that once a receiver adopts a product, each sender who talked to the receiver experiences the same externalities of r. That is, if there are m senders for a given receiver, then the total externalities generated by the receiver are mr. The reward can be conditioned on the set of senders who talked. We assume that G follows (4).

Let $m > 1$ be the number of senders connected to a given receiver. Suppose that, when there is only one sender, R is the optimal expected referral reward. The conclusion in Lemma 1 or the analysis in Section 4.5 entails that, by paying R in expectation to each sender, the firm can give the same incentive of talking to the senders. However, such an adjustment changes the firm's total payment. This is because, the expected payment of referral reward is no longer R, but mR.

This implies that the firm becomes reluctant to use referral rewards. More precisely, if the optimal reward level is zero in the model with one sender per receiver, then it is still zero in the model with multiple senders per receiver. At the same time, free contracts become relatively more attractive as it incentivizes senders in the same way as with only one sender. Thus, when there are multiple senders per receiver, the range of parameter values such that only free contracts are used becomes wider because free contracts can substitute referral rewards.

4.7 Social Optimum

In order to understand the monopolist's strategy better, we consider the social planner's solution and compare it with the solution obtained in the main section. Specifically, we consider a social planner who has control over the senders' actions $a_i \in \{\text{Refer}, \text{Not}\}$ and the quantities q_L and q_H offered to receivers, while she does not have control over receivers' choice of whether to actually use the product after it is allocated.[24] Rewards and prices do not show up in the social planner's problem because they are only transfers between agents.

We start with two basic observations. First, whenever WoM takes place under the monopolist's solution, there is a surplus from WoM. Hence, it is also in the social planner's interest to encourage WoM. Second, under the monopolist's optimal scheme, free contracts

[24] In the classic setup of Maskin and Riley (1984), all buyers get positive utility from using the product, and thus, they always use the product after purchase. If we were to allow the social planner to have control over the use of the product and $v'_L(q) < c$ for all $q > 0$, then she would have low types use just a little bit of the quantity and generate the externalities r, which we view as implausible.

always make senders weakly better off by increasing the probability of receiving externalities, high-type receivers better off by reducing the price due to the information rent, and low-type receivers indifferent because their participation constraint is always binding. This implies that, if the monopolist firm optimally offers free contracts, then it is also socially optimal to offer it. We summarize these two observations in the following proposition:

Proposition 6. *1. If there exists a monopolist's solution under which $a_i = Refer$ for all i, then there exists a social planner's solution that entails $a_i = Refer$ for all i.*

2. If there exists $((0, \underline{q}), (\tilde{p}^*_H, q^*_H), R) \in \mathcal{S}$ for some R under the monopolist's solution, then there exists a social planner's solution that entails $q_L = \underline{q}$.

The converse of each part of the above proposition is not necessarily true, i.e., the monopolist may be less willing to encourage WoM than the social planner or not offer free contracts despite it being socially optimal. To see this clearly, we further investigate the social planner's problem in what follows.

Conditional on free contracts being offered, the welfare-maximizing menu of quantities (q_H, q_L) is exactly the same as the menu offered by the monopolist in the main section. To see why, first note that, as in the classic screening problem in Maskin and Riley (1984), the monopolist's solution results in no distortions at the top, i.e., $v'(q_H) = c$. Conditional on selling to the low types, the low-type quantity q_L under the second best in Maskin and Riley (1984) is distorted to deter high types to switch to the contract offered to low types. This means that the social planner's solution dictates that low types receive more quantity in the first best than in the second best. In our problem, however, the welfare-maximizing quantity cannot be strictly higher than $\underline{q}$ because the marginal cost c is higher than the marginal benefit $v'_L(q)$ for all $q \geq \underline{q}$ (Assumption 2), and the incentive-compatible quantity cannot be strictly lower than $\underline{q}$ because the low types would not use the product for $q_L < \underline{q}$.

Finally, whether or not the sender talks under the social planner's solution depends on the comparison between the total benefit from talking and the cost of talking, $\bar{\xi}$: In total,

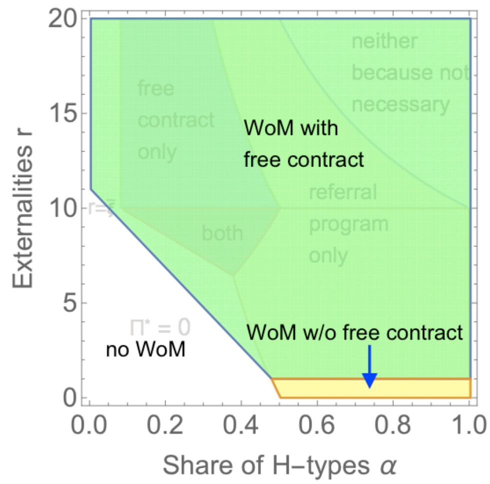


Figure 4: Socially optimal WoM in the (α, r)-space: Under the social optimum, (i) the senders engage in WoM if and only if the parameters fall in the colored parameter region, and (ii) free contracts are used if and only if the parameters fall in the top-right region tagged as "free contract." The background displays the monopolist's solution as presented in Figure 2.

WoM is efficient if and only if

$$\alpha(v_H(q_H^*) - cq_H^* + r) + (1-\alpha)\max\{r - c\underline{q}, 0\} \geq \bar{\xi}. \tag{12}$$

Note that there are two social benefits of WoM. First, WoM creates network externalities because the senders and receivers become aware of each other using the product. Second, it creates gains from trade because some high-valuation buyers learn about the product. Figure 4 summarizes the above findings using the same parameters as in Figure 2.

In the monopolist's solution, free contracts are not used if $r < \frac{CF^*}{1-\alpha}$. Substituting the definition of CF^*, shows that this is equivalent to $r - c\underline{q} < \frac{\alpha}{1-\alpha} v_H(\underline{q})$. Since the social planner uses free contracts if $0 < r - c\underline{q}$, the monopolist uses free contracts too little from the social planner's point of view conditional on it being socially optimal to encourage WoM if r is high, and α or $v_H(\underline{q})$ is high. The reason is as follows. On the one hand, high externalities r imply a high additional benefit r from having a receiver using the product, so that the social planner wants all receivers to use the product. However, such r pertains to the senders and the monopolist cannot extract the entire corresponding surplus. On the other hand, the monopolist is reluctant to use free contracts if the information rent necessary to induce high types to purchase a premium contracts is high relative to the number of low types who choose the free contracts. The "per low-type" information rent $\frac{\alpha}{1-\alpha} v_H(\underline{q})$ is high if α or $v_H(\underline{q})$ is high.

5 Conclusion

The case of Dropbox shows that WoM plays an important role in customer acquisition. This paper is the first to incorporate WoM in a contracting problem. We jointly analyzed the role of a freemium strategy and referral rewards when incentivizing WoM for products with positive externalities.

We present a model of optimal contracting in which the number of customers depends on WoM. The monopolist firm optimally encourages senders of the information to engage in WoM by fine-tuning two parts of the benefit of talking: referral rewards and expected externalities.

Despite being very simple, the model allows for a rich set of predictions. Depending on the environment, it is optimal to use one, both or none of these methods. We show that it is optimal to use referral programs when the size of externalities is small, and free versions are useful when there are many low-type customers. The pattern of the optimal scheme is consistent with the strategies we observe for companies such as Dropbox, Skype, Uber, and Amtrak.

We keep our model particularly simple and there are many ways to enrich it. We have enumerated potential reasons for the use of free products in the Related Literature, and it would be interesting to build a model that includes those effects as well. In such extensions, the findings in this paper would be helpful in identifying the implication of the those additional effects. One feature of our model that may be unrealistic is that low-type receivers enjoy zero surplus. Such a feature would disappear once we have more than two types (such as a continuous type model). Even in such a model, the basic feature of the optimal scheme would be the same; for example, under certain parameter values free contracts would be purchased with positive probability as long as there are some types whose valuations are lower than the cost of production.

Another extension of interest is the one in which receivers are uncertain about the quality of the product, and the senders have higher incentives to talk when the he knows the quality is higher. In such a model, if the receivers know that the senders would receive referral rewards, then they may adjust their belief about the quality downwards. Although this may be a worthwhile direction to extend the model, it would require too much divergence from the Maskin-Riley model, and hence is outside the scope of the current paper.

Finally, our work suggests possibilities of empirical research. It may help estimate the value of externalities that the senders perceive upon referring. We hope our paper stimulates a sequence of such research.

References

Ajorlou, Amir, Ali Jadbabaie, and Ali Kakhbod, "Dynamic Pricing in Social Networks: The Word of Mouth Effect," Technical Report, Mimeo 2015.

Berger, Jonah, "Word of mouth and interpersonal communication: A review and directions for future research," *Journal of Consumer Psychology*, 2014, *24* (4), 586–607.

_ **and Eric M Schwartz**, "What drives immediate and ongoing word of mouth?," *Journal of Marketing Research*, 2011, *48* (5), 869–880.

Biyalogorsky, Eyal, Eitan Gerstner, and Barak Libai, "Customer referral management: Optimal reward programs," *Marketing Science*, 2001, *20* (1), 82–95.

Campbell, Arthur, "Word of mouth and percolation in social networks," *American Economic Review*, 2012.

_ , Dina Mayzlin, and Jiwoong Shin, "Managing Buzz," 2015.

Csorba, Gergely, "Screening contracts in the presence of positive network effects," *International Journal of Industrial Organization*, 2008, *26* (1), 213–226.

Dewatripont, Mathias, "Costly communication and incentives," *Journal of the European Economic Association*, 2006, *4* (2-3), 253–268.

_ and Jean Tirole, "Modes of communication," *Journal of Political Economy*, 2005, *113* (6), 1217–1238.

Galeotti, Andrea, "Talking, searching, and pricing*," *International Economic Review*, 2010, *51* (4), 1159–1174.

_ and Sanjeev Goyal, "Influencing the influencers: a theory of strategic diffusion," *The RAND Journal of Economics*, 2009, *40* (3), 509–532.

Godes, David and Dina Mayzlin, "Firm-created word-of-mouth communication: Evidence from a field test," *Marketing Science*, 2009, *28* (4), 721–739.

_ , _ , Yubo Chen, Sanjiv Das, Chrysanthos Dellarocas, Bruce Pfeiffer, Barak Libai, Subrata Sen, Mengze Shi, and Peeter Verlegh, "The firm's management of social interactions," *Marketing Letters*, 2005, *16* (3-4), 415–428.

Hahn, Jong-Hee, "Nonlinear pricing of telecommunications with call and network externalities," *International Journal of Industrial Organization*, 2003, *21* (7), 949–967.

Houston, Drew, 2010. `http://www.slideshare.net/gueste94e4c/dropbox-startup-lessons-learned-3836587`. Accessed: 2016-05-05.

Lee, Clarence, Vineet Kumar, and Sunil Gupta, "Designing freemium: a model of consumer usage, upgrade, and referral dynamics," Technical Report, Mimeo 2013.

_ , _ , and _ , "Designing Freemium: Balancing Growth and Monetization Strategies," 2015.

Maskin, Eric and John Riley, "Monopoly with incomplete information," *The RAND Journal of Economics*, 1984, *15* (2), 171–196.

Segal, Ilya, "Coordination and discrimination in contracting with externalities: Divide and conquer?," *Journal of Economic Theory*, 2003, *113* (2), 147–181.

Shapiro, Carl and Hal R Varian, "Versioning: the smart way to," *Harvard Business Review*, 1998, *107* (6), 107.

Shi, Zijun, Kannan Srinivasan, and Kaifu Zhang, "Product Line Design under Network Externalities: On the Optimality of Freemium," Technical Report, Mimeo 2015.

Economist

Economist, 2012. `http://www.economist.com/blogs/babbage/2012/12/dropbox`. Accessed: 2016-06-27.

UBER Newsroom, 2014. `https://newsroom.uber.com/were-doubling-referral-credits-for-the-new-year-2/`. Accessed: 2016-05-05.

Woirhaye, Nicolas, "LinkedIn against its European competitors (A comparative study of LinkedIn vs Viadeo vs Xing)," *IKO System*, 2011. `http://www.iko-system.com/wp-content/uploads/2014/02/LinkedIn-vs-competitors.pdf`. Accessed: 2016-06-27.

A Appendix: Proofs

Proof. **(Proposition 1)** First, we show that it is without loss of generality to restrict attention to choice variables in a compact set. To see this, first note that, as we will show in the proof of Lemma 2, a scheme $((p_L, q_L), (p_H, q_H), R)$ with $q_L \in (0, \underline{q})$ generates a strictly lower profit than a scheme $((p_L, 0), (p_H, q_H), R)$. The same proof also shows that a scheme $((p_L, q_L), (p_H, q_H), R)$ with $q_L > \underline{q}$ generates a strictly lower profit than a scheme $((p_L, \underline{q}), (p_H, q_H), R)$. Thus it is without loss of generality to restrict attention to $\{0, \underline{q}\}$ as the space from which q_L is chosen. This and the the participation constraint for low types implies that if a scheme $((p_L, q_L), (p_H, q_H), R)$ satisfies the constraints then $p_L \leq 0$. Also, the proof for Lemma 2 shows that for any scheme $((p_L, q_L), (p_H, q_H), R)$, $p_L < 0$ implies that the participation constraints for both types are non-binding, hence there exists $\epsilon > 0$ such that there exists a scheme $((p_L + \epsilon, q_L), (p_H + \epsilon, q_H), R)$ that satisfies the constraints and generates a higher profit than the original scheme. Consequently, it is without loss of generality to restrict attention to a scheme $((p_L, q_L), (p_H, q_H), R)$ with $p_L = 0$.

Also, since $\lim_{q \to \infty} v_H'(q) < c$, there exists q' such that any scheme $((p_L, q_L), (p_H, q_H), R)$ with $q_H > q'$ generates a strictly negative profit. Thus it is without loss of generality to restrict attention to $[0, q']$ for the space for q_H, where q' is any number satisfying $v_H'(q') < c$. Fix such $q' < \infty$ arbitrarily. Then, any scheme $((p_L, q_L), (p_H, q_H), R)$ with $R > v_H(q')$ generates a strictly negative profit, so again it is without loss to restrict attention to $[0, v_H(q')]$ as the space for R.

These bounds for q_H and q_L together with the PC constraints imply that it is without loss of generality to consider $p_H \leq v_H(q')$. The incentive compatibility condition for low types implies that $0 = \max\{v_L(q_L), 0\} - p_L \geq \max\{v_L(q_H), 0\} - p_H$, which implies $p_H \geq \max\{v_L(q_H), 0\} \geq 0$. Thus, it is without loss of generality to consider $p_H \in [0, v_H(q')]$.

These facts and the fact that all constraints are weak inequalities with continuous functions imply that the optimal scheme is chosen from a compact set. Now, note that the objective function is right-continuous in each choice variable because G is a cumulative distribution function, and all jumps are upwards.

These facts and the assumption that G has only finitely many discontinuities imply that there exists a partition of the compact space of the choice variables C with a finite number of cells $(P_1, \dots, P_K)$ for some integer $K \in \mathbb{N}$, such that over each cell, the objective function is continuous.

Let $\hat{\pi}$ be the supremum of the objective function over C. Then there exists a sequence $(y^k)_{k=1,2,\dots}$ with $y^k \in C$ for all k such that the value of the objective function under y^k converges to $\hat{\pi}$. Since $K < \infty$, this implies that there exists a cell of the partition, denoted P_{i^*} (choose one arbitrarily if there are multiples of such cells), and a subsequence $(z^k)_{k=1,2,\dots}$ of $(y^k)_{k=1,2,\dots}$ such that $z^k \in P_{i^*}$ for all k.

Since P_{i^*} is a bounded set, $(z^k)_{k=1,2,\dots}$ has an accumulation point. Let an arbitrary choice of an accumulation point be z^*. If $z^* \in P_{i^*}$, then by continuity the objective function attains the value $\hat{\pi}$ at z^*. If $z^* \notin P_{i^*}$, then by the assumption of the upward jumps, the objective function attains the value strictly greater than $\hat{\pi}$ at z^*, which is a contradiction. This completes the proof. □

Proof. **(Lemma 2)** Let $((p_L, q_L), (p_H, q_H), R)$ be an optimal scheme.

(i) Given a menu of contracts with $q_L > \underline{q}$ that satisfy (2), continuity of v_L implies that the monopolist can decrease q_L and p_L slightly, such that $\max\{v_L(q_L), 0\} - p_L$ remains constant (by Assumption 1) without violating (2) because $v_H(q_L) - p_L$ decreases with such a change (as $v'_H > v'_L$). This strictly increases profits by Assumption 2. Similarly, given a menu of contracts with $0 < q_L < \underline{q}$ that satisfy (2) and such that $\Pi^* > 0$, the monopolist can decrease q_L to zero and increase profits without violating (2).

The equation $p_L = 0$ can be shown by noting that type L's participation constraint must be binding: Assume $p_L < \max\{v_L(q_L), 0\} = 0$. First, note that then type H's participation constraint cannot be binding: If it was, then

$$0 = \max\{v_H(q_H), 0\} - p_H \geq \max\{v_H(q_L), 0\} - p_L \geq \max\{v_L(q_L), 0\} - p_L > 0$$

which is a contradiction. Thus, the monopolist can strictly increase profits by increasing p_L and p_H by the same small amount such that (2) remains to be satisfied. Consequently,

$p_L = \max\{v_L(q_L), 0\} = 0$.

(ii) Given a R, $p_L = 0$ and fixing $q_L \in \{0, \underline{q}\}$, H-type's contract (p_H, q_H) must solve $\max_{p_H, q_H} \alpha(p_H - q_H c)$ subject to $\max\{v_H(q_H), 0\} - p_H \geq \max\{v_H(q_L), 0\}$ and $\max\{v_H(q_H), 0\} - p_H \geq 0$. If we ignored the participation constraint, and solved a relaxed problem, the incentive compatibility constraint must be binding and it follows that $q_H = q_H^*$ and $p_H = \max\{v_H(q_H^*), 0\} - \max\{v_H(q_L), 0\}$. This automatically satisfies the participation constraint:

$$\max\{v_H(q_H^*), 0\} - [\max\{v_H(q_H^*), 0\} - \max\{v_H(q_L), 0\}] = \max\{v_H(q_L), 0\} > \max\{v_L(q_L), 0\} = 0.$$

The above proof shows that IC constraint of the H-type is binding. Using this fact, parts (iii) and (iv) follow by plugging q_L into type-H's incentive compatibility constraint. □

Proof. **(Lemma 3, Referral Program)** A sender talks if and only if

$$\overline{\xi} \leq r \left(\alpha + (1 - \alpha) \cdot \mathbf{1}_{\{q_L > 0, v_L(q_L) \geq 0\}}\right) + R.$$

As a result, the monopolist must pay at least (5) in order to assure that senders talk and thus, the monopolist pays exactly this as long as it is profitable to inform receivers, i.e., as long as $R^{**}((p_L, q_L), (p_H, q_H)) < \pi((p_L, q_L), (p_H, q_H))$ holds. □

Proof. **(Lemma 4, Free Contracts)** (i) If $\overline{\xi} \leq \alpha r$, then the senders' IC constraint is always satisfied, so that the seller's problem collapses to

$$\max_{p_L, p_H \in \mathbb{R},\, q_L, q_H \geq 0} N \cdot \left[\alpha \cdot (p_H - q_H \cdot c) + (1 - \alpha) \cdot (p_L - q_L \cdot c) - R\right]$$

which is equivalent to the maximization problem in the benchmark case with free WoM. Thus, no free contracts are offered under any optimal scheme.

(ii) First, note that if $\Pi^* > 0$, it suffices to show when profits with free contracts (and the optimal reward scheme given by Lemma 3) are greater than profits without free contracts.

Let $\alpha r < \overline{\xi} \leq r$. First, if $\overline{\xi} - \alpha r > \Pi^{\text{classic}}$, then by Lemma 3 not offering free contracts yields negative profits and cannot be optimal. If $\overline{\xi} - \alpha r \leq \Pi^{\text{classic}}$, then by Lemma 3, the optimal reward is $R = 0$ whenever $q_L = \underline{q}$ and is $R = \overline{\xi} - \alpha r$ whenever $q_L = 0$. With $p_L = 0$ and (p_H, q_H) as in Lemma 2, it follows immediately that offering free contracts generates weakly higher profits than offering $q_L = 0$ if and only if $\Pi^{\text{classic}} - \alpha v_H(\underline{q}) - (1-\alpha) \cdot \underline{q} \cdot c \geq \Pi^{\text{classic}} - (\overline{\xi} - \alpha r)$, which is equivalent to (8).

(iii) Let $\overline{\xi} > r$. Then, by Lemma 3 if the monopolist chooses $q_L = \underline{q}$, then profits are given by $\Pi^{\text{classic}} - CF^* - (\overline{\xi} - r)$ and if $q_L = 0$, then profits are given by $\Pi^{\text{classic}} - (\overline{\xi} - \alpha r)$. Thus, offering free contracts generates a weakly higher profit than offering no free contracts if and only if $\Pi^{\text{classic}} - CF^* - (\overline{\xi} - r) \geq \Pi^{\text{classic}} - (\overline{\xi} - \alpha r)$, which is equivalent to $CF^* \leq (1-\alpha)r$. □

Proof. **(Theorem 1, Full Characterization)** 1. By Lemmas 2 and 3, $\Pi^* > 0$ if and only if $\Pi^{\text{classic}} - CF^* - \max\{\overline{\xi} - r, 0\} > 0$ or $\Pi^{\text{classic}} - \max\{\overline{\xi} - \alpha r, 0\} > 0$. Since $\Pi^{\text{classic}} > 0$, this can be rewritten as $\Pi^{\text{classic}} - CF^* - \max\{\overline{\xi} - r, 0\} > 0$ or $\Pi^{\text{classic}} - (\overline{\xi} - \alpha r) > 0$.

2. This follows immediately from Lemma 4.

3. (a) By Lemma 3, in the presence of free contracts, a reward must only be paid if $r > \overline{\xi}$.

(b) Similarly, if no free contracts are offered, positive rewards are only being paid if $\alpha r < \overline{\xi}$. □

Proof. **(Proposition 2)** (i) Denote the maximal expected profit without free contracts (i.e., $q_L = 0$ is offered to low types) under α by $\Pi^{\text{not free}}(\alpha)$. Similarly, denote the maximal expected profit with free contracts under α by $\Pi^{\text{free}}(\alpha)$.[25] The function $\Pi^{\text{not free}}(\alpha)$ is concave as long as $\Pi^{\text{not free}}(\alpha) > 0$, and $\Pi^{\text{free}}(\alpha)$ is linear in α as long as $\Pi^{\text{free}}(\alpha) > 0$. Moreover, we have that

$$\begin{aligned} \lim_{\alpha \to 1} \Pi^{\text{free}}(\alpha) &= \lim_{\alpha \to 1} \alpha(p_H^* - q_H^* c - v_H(\underline{q})) - (1-\alpha)\underline{q}c - \max\{\xi - r, 0\} \\ &< \lim_{\alpha \to 1} \alpha(p_H^* - q_H^* c) - \max\{\xi - \alpha r, 0\} = \Pi^{\text{not free}}(\alpha). \end{aligned}$$

[25] Existence of these maxima follows from an analogous proof to the one for Proposition 1.

This implies that $\Pi^{\text{not free}}(\alpha)$ and $\Pi^{\text{free}}(\alpha)$ intersect at most once. Hence, if $\Pi^{\text{free}}(\alpha_1) \geq \Pi^{\text{not free}}(\alpha_1)$, then $\Pi^{\text{free}}(\alpha_2) > \Pi^{\text{not free}}(\alpha_2)$ for all $\alpha_2 < \alpha_1$. This concludes the proof.

(ii) This part follows directly from part 2 of Theorem 1. □

Proof. **(Proposition 3)** By Lemma 3, we have

$$R^{**}((0,0),(p_H^1,q_H^1)) \;=\; \max\{\overline{\xi}-\alpha r,\ 0\} > \max\{\overline{\xi}-r,\ 0\} \geq R^{**}((0,\underline{q}),(p_H^2,q_H^2))$$

because $R^{**}((0,0),(p_H^1,q_H^1)) < \pi((0,0),(p_H^1,q_H^1))$ and $\overline{\xi}-\alpha r > 0$. □

Proof. **(Proposition 4)** 1. If all senders choose Refer, the IC constraints for all senders—those who see s_H and those who see s_L— must be satisfied. (a) Without free contracts, the senders' IC constraints are given by:

$$\overline{\xi} \leq \alpha_H r + (\alpha_H \mathbf{R}(H) + (1-\alpha_H)\mathbf{R}(L)) \quad \text{and} \quad \overline{\xi} \leq \alpha_L r + (\alpha_L \mathbf{R}(H) + (1-\alpha_L)\mathbf{R}(L)).$$

The optimal reward conditional on these constraints minimizes referral reward payments by making both senders' IC constraints binding whenever possible. The firm is able to do this if and only if $r \leq \overline{\xi}$ and in that case the optimal reward scheme is given by $\mathbf{R}(H) = \overline{\xi} - r$ and $\mathbf{R}(L) = \overline{\xi}$. If $r > \overline{\xi}$, it is optimal to set $\mathbf{R}(H) = 0$ and $\mathbf{R}(L) = \max\left\{\frac{\overline{\xi}-\alpha_L r}{1-\alpha_L}, 0\right\}$.

(b) With free contracts, the senders' IC constraints are given by:

$$\overline{\xi} \leq r + (\alpha_H \mathbf{R}(H) + (1-\alpha_H)\mathbf{R}(L)) \quad \text{and} \quad \overline{\xi} \leq r + (\alpha_L \mathbf{R}(H) + (1-\alpha_L)\mathbf{R}(L)).$$

Thus, it is optimal to set $\mathbf{R}(H) = \mathbf{R}(L) = \max\{\overline{\xi} - r, 0\}$.

2. If senders who saw s_L do not talk, then only the IC constraint of a sender who sees s_H must be satisfied and the IC constraint of the sender who sees s_L must be violated.

(a) Without free contracts, the firm minimizes reward payments subject to these constraints by minimizing $\alpha_H \mathbf{R}(H) + (1-\alpha_H)\mathbf{R}(L)$ (i.e., making the IC for the sender with s_H

binding whenever possible) such that

$$\alpha_L r + (\alpha_L \mathbf{R}(H) + (1-\alpha_L)\mathbf{R}(L)) < \overline{\xi} \leq \alpha_H r + (\alpha_H \mathbf{R}(H) + (1-\alpha_H)\mathbf{R}(L)).$$

First, note that these inequalities imply $\mathbf{R}(H) > \mathbf{R}(L) - r$. Second, if a referral scheme with $\mathbf{R}(H), \mathbf{R}(L) \geq 0$ that satisfies these inequalities exists (this is the case whenever $\frac{\overline{\xi}}{\alpha_L} - r \geq 0$), then the referral scheme given by $\mathbf{R}(L) = 0$, $\mathbf{R}(H) = \max\{\frac{\overline{\xi}}{\alpha_H} - r, 0\}$ must maximize the seller's profits: The seller cannot increase profits by decreasing $\alpha_H \mathbf{R}(H) + (1-\alpha_H)\mathbf{R}(L)$.

(b) With free contracts, the constraints become

$$r + (\alpha_L \mathbf{R}(H) + (1-\alpha_L)\mathbf{R}(L)) < \overline{\xi} \leq r + (\alpha_H \mathbf{R}(H) + (1-\alpha_H)\mathbf{R}(L)),$$

which imply $\mathbf{R}(H) > \mathbf{R}(L)$. By an analogous argument as in (a), a reward scheme satisfying these constraints exists if and only if $\overline{\xi} - r \geq 0$ and in that case the scheme given by $\mathbf{R}(H) = \frac{\overline{\xi}-r}{\alpha_H}$, $\mathbf{R}(L) = 0$ maximizes profits. □

Proof. **(Proposition 5)** 1. First, note that any optimal scheme results in one of the following three types of behaviors by the senders: Either (i) no senders talks, or (ii) all senders talk, or (iii) only senders who have received a s_H signal talk.[26]

If $\overline{\xi} - r \geq \alpha(p_H^* - cq_H^*)$, then for all $\beta \in (\frac{1}{2}, 1)$ the firm cannot make positive profits. We assume from now on $\overline{\xi} - r < \alpha(p_H^* - cq_H^*)$. We will show that for sufficiently large β, the firm can make positive profits, i.e., that we are in case (ii) or (iii).

Fix $\beta \in (\frac{1}{2}, 1)$. If $\overline{\xi} - r\alpha_L \leq 0$, then all senders talk even without any reward payments as long as H-type receivers consume a positive quantity. Thus, we are in case (ii), and so for any optimal scheme $((p_H, q_H), (p_L, q_L), \mathbf{R})$, $\mathbf{R}(L) = 0$ and $q_L = 0$ hold.

We assume from now on that $r\alpha_L < \overline{\xi} < \alpha(p_H^* - cq_H^*) + r$. Under a reward scheme $\mathbf{R}$

[26]Note that there is no optimal scheme in which s_L-senders talk while s_H-senders do not talk. This is because $\alpha_H > \alpha_L$ and thus, given a scheme $((p_H, q_H), (p_L, q_L), \mathbf{R})$ where only s_L-senders talk, the seller can strictly increase profits by choosing a reward scheme $\mathbf{R}'$ with $\mathbf{R}'(H) = \mathbf{R}'(L) = \alpha_L \mathbf{R}(H) + (1-\alpha_L)\mathbf{R}(L)$ while holding the menu of contracts fixed. Under this scheme, both sender types talk.

with $\mathbf{R}(L) = 0$ (as specified in Proposition 4) and $\mathbf{R}(H) = \frac{\max\{\bar{\xi}-\alpha_H r, 0\}}{\alpha_H}$, the senders who have seen s_H talk, while senders who have seen s_L do not talk.

Next we show that, there exists $\bar{\beta} < 1$ such that for all $\beta > \bar{\beta}$, it is not optimal to offer free contracts and the firm always chooses to be in case (iii). For this purpose, we compute the profits from cases (ii) and (iii).

- **Case (iii):** Since $\alpha_H \to 1$ as $\beta \to 1$, there exists $\bar{\beta} < 1$ such that for all $\beta > \bar{\beta}$, it is not optimal to offer free contracts by the analysis in Section 3. Thus, the profits are given by $\alpha\beta(p_H^* - cq_H^*) - (\alpha\beta + (1-\alpha)(1-\beta)) \max\{\bar{\xi} - \alpha_H r, 0\}$, which is greater than zero for sufficiently large β because it converges to $\overline{\Pi}_H^* \equiv \alpha(p_H^* - cq_H^*) - \alpha \max\{\bar{\xi} - r, 0\} \geq \max\{\alpha(p_H^* - cq_H^*) - (\bar{\xi} - r), \alpha(p_H^* - cq_H^*)\} > 0$ as $\beta \to 1$.

- **Case (ii):** We consider two cases: $\bar{\xi} \geq r$ and $\bar{\xi} < r$.

  - $\bar{\xi} \geq r$**:** By Proposition 4, without free contracts, profits are given by $\alpha(p_H^* - cq_H^*) - (\bar{\xi} - \alpha r)$ and with free contracts they are given by $\alpha(p_H^* - cq_H^*) - CF^* - (\bar{\xi} - r)$. Both profits are strictly smaller than $\overline{\Pi}_H^*$.
  - $\bar{\xi} < r$**:** Without free contracts, profits are given by $\alpha(p_H^* - cq_H^*) - (1-\alpha) \max\left\{\frac{\bar{\xi} - \alpha_L r}{1-\alpha_L}, 0\right\}$ and with free contracts, they are $\alpha(p_H^* - cq_H^*) - CF^*$. Both profits converge to numbers that are smaller than $\overline{\Pi}_H^*$ as $\beta \to 1$.

Hence, there exists $\bar{\beta} < 1$ such that for all $\beta > \bar{\beta}$, it is not optimal to offer free contracts and the firm always chooses to be in case (iii). This concludes the proof.

2. If $\beta = \frac{1}{2}$, then one can immediately see from the expressions above that profits coincide with the ones in the main section. Thus, by continuity, for any $r < \frac{\bar{\xi}}{\alpha}$, there exists a $\bar{\beta} > \frac{1}{2}$ such that for all $\beta \in (\frac{1}{2}, \bar{\beta})$, $r < \frac{\bar{\xi}}{\alpha_L}$ and $r < \frac{\bar{\xi}}{\alpha_H}$. Similarly, for any $r \in \left(\frac{\bar{\xi}}{\alpha_L}, \frac{CF^*}{1-\alpha_L}\right)$, there exists a $\bar{\beta} > \frac{1}{2}$ such that for all $\beta \in (\frac{1}{2}, \bar{\beta})$, $r \in \left(\frac{\bar{\xi}}{\alpha_L}, \frac{CF^*}{1-\alpha_L}\right)$ and $r \in \left(\frac{\bar{\xi}}{\alpha_H}, \frac{CF^*}{1-\alpha_H}\right)$. Analogous conclusions hold for intervals $\left(\frac{CF^*}{1-\alpha}, \frac{\bar{\xi}-CF^*}{\alpha}\right)$ and $\left(\frac{\bar{\xi}-CF^*}{\alpha}, \infty\right)$. Thus, there exists a $\bar{\beta} > \frac{1}{2}$ such that for all $\beta \in (\frac{1}{2}, \bar{\beta})$, the same analysis as in the main section applies for β. $\square$

Online Supplementary Appendix to: Encouraging Word of Mouth: Free Contracts, Referral Programs, or Both?

Yuichiro Kamada and Aniko Öry*

July 12, 2016

B Heterogeneous Costs of WoM

In this Online Appendix, we consider the case with heterogeneous costs of talking. We restrict attention to twice differentiable G with $G' = g$ satisfying $g(\xi) > 0$ for all $\xi \in \mathbb{R}_+$ and

Assumption 4. *G is strictly log-concave, i.e., $\frac{g}{G}$ is strictly decreasing.*

This condition is satisfied by a wide range of distributions such as exponential distributions, a class of gamma, Weibull, and chi-square distributions, among others.

Section B.1 characterizes the optimal scheme. Section B.2 conducts comparative statics of the optimal scheme. Section B.3 contains all the proofs for these results. Section B.4 discusses how the main model with homogeneous costs can be viewed as a limit of models with heterogeneous costs.

B.1 Properties of Optimal Contracts

First, we characterize the optimal reward. If free contracts are offered, it acts as a substitute for reward payments, which results in higher optimal rewards absent free contracts. The following proposition provides conditions under which a positive reward is optimally offered.

Lemma 5 (Optimal Reward)**.** *In the model with heterogeneous costs, there exists r^{free} and $r^{not\ free}$ with $r^{not\ free} > r^{free}$ such that the following are true:*

*Kamada: Haas School of Business, University of California, Berkeley, Berkeley, CA 94720, e-mail: y.cam.24@gmail.com; Öry: School of Management, Yale University, New Haven, CT 06511, e-mail: aniko.oery@yale.edu.

1. *If* $r < r^{free}$*, then* $((p_L, q_L), (p_H, q_H), R) \in \mathcal{S}$ *implies* $R > 0$.

2. *If* $r^{free} \leq r < r^{not\ free}$*, then* $((p_L, q_L), (p_H, q_H), R) \in \mathcal{S}$ *implies either* $R > 0$ *and* $q_L = 0$*, or* $R = 0$ *and* $q_L = \underline{q}$.

3. *If* $r^{not\ free} \leq r$*, then* $((p_L, q_L), (p_H, q_H), R) \in \mathcal{S}$ *implies* $R = 0$.

In order to prove this, we fix a menu of contracts with and without free contracts satisfying the conditions in Lemma 2 and solve for the optimal reward scheme. That is, conditional on offering free contracts ($q_L = \underline{q}$), we define the maximal profit under (r, α) by

$$\Pi^{\text{free}}(r, \alpha) = \max_{R \geq 0} \left(\left[\pi((0, \underline{q}), (\tilde{p}_H^*, q_H^*)) - R \right] \cdot G(r + R) \right)$$

and conditional on offering no free contracts ($q_L = 0$), define the maximal profit under (r, α) by

$$\Pi^{\text{not free}}(r, \alpha) = \max_{R \geq 0} \left([\pi((0, 0), (p_H^*, q_H^*)) - R] \cdot G(\alpha r + R) \right).$$

Let us also define the unique optimal reward given that free contracts are offered and that no free contracts are offered by $R^{\text{free}}(r, \alpha)$ and $R^{\text{not free}}(r, \alpha)$, respectively.

There are three reasons why $r^{\text{not free}} > r^{\text{free}}$ holds. As opposed to a situation without free contracts, with free contracts, (i) positive quantity is offered to low types, (ii) information rent is provided to high types, and (iii) the sender receives full externalities conditional on talking. All these effects reduce the incentive to provide referral rewards. Note that $r^{\text{not free}}$ corresponds to $\frac{\overline{\xi}}{\alpha}$ in the homogeneous model, while r^{free} corresponds to $\overline{\xi}$. In the homogeneous-cost setting, only reason (iii) affected the comparison of r^{free} and $r^{\text{not free}}$. The effects (i) and (ii) were present, but they only determined whether offering free contracts generates nonnegative profits.

The following theorem summarizes some general properties of optimal contracts. Unlike Theorem 1, it is not a full characterization, but it shows that many features of the optimal scheme with homogeneous cost carries over to the ones for heterogeneous costs.

Theorem 2 (Optimal Contracts)**.** *The following claims hold in the model with heterogeneous costs:*

1. **(Positive profits)** $\Pi^{not\, free}(r,\alpha) > 0$ *for all* $r \in [0,\infty)$ *and* $\alpha \in (0,1)$.

2. **(Using both rewards and free contracts)** *There exists* $((0,\underline{q}),(\tilde{p}^*_H, q^*_H), R) \in \mathcal{S}$ *such that* $R > 0$ *(i.e., it is optimal to provide both free contracts and rewards) if and only if*

$$r^{free} > r \geq \frac{CF^*}{1-\alpha}. \tag{13}$$

3. *Suppose that* $\frac{G(\xi)}{g(\xi)}$ *is convex.*

   (a) **(Free vs. no free contracts)** *There exist* $\underline{r}, \overline{r} \in [\frac{CF^*}{1-\alpha}, \infty)$ *such that there exists* $((0,\underline{q}),(\tilde{p}^*_H, q^*_H), R) \in \mathcal{S}$ *for some* $R \geq 0$ *(i.e., it is optimal to provide free contracts) if and only if* $r \in [\underline{r}, \overline{r}]$.

   (b) **(Never free contracts)** *If* $\frac{CF^*}{1-\alpha} > r^{not\, free}$, *then* $[\underline{r}, \overline{r}] = \emptyset$.

First, unlike in the homogeneous-cost model, profits without offering free contracts are always positive: With homogeneous costs, profits without free contracts are negative when the share of high types are low, so the expected externalities are low. This is because low expected externalities imply that a sufficient size of reward is necessary to encourage WoM, but such a cost cannot be compensated by the profits generated by only a small fraction of high types. With heterogenous costs, there always exists some fraction of customers with sufficiently small WoM costs, who do not need to be rewarded to initiate referrals.

Part 2 of the proposition shows that even with heterogeneous costs we can derive necessary and sufficient conditions for a combination of free contracts and rewards programs to be offered. As with homogeneous cost, free contracts are only optimal for sufficiently large externalities r and rewards are only offered for sufficiently small externalities.

Externalities	$r < r^{\text{free}}$	$r^{\text{free}} < r < r^{\text{not free}}$			$r^{\text{not free}} < r$
Referral rewards	Yes	No	or	Yes	No
Free contracts	No $\Leftrightarrow r < \frac{CF^*}{1-\alpha}$	Yes		No	Yes $\Leftrightarrow r$ is small

Table 2: Comparative Statics with respect to r with heterogeneous WoM costs

For a full characterization of the optimal menu of contracts, it is useful to impose the additional assumption that $\frac{G}{g}$ is convex. This condition is, for example, satisfied by the exponential distribution. Given this assumption, free contracts are only offered for an intermediate connected range of externalities r. We can extend these results qualitatively as follows.

Remark 1. If we do not impose $\frac{G}{g}$ to be convex, one can still show that $\lim_{r\to 0} \Pi^{\text{not free}}(r, \alpha) > \lim_{r\to 0} \Pi^{\text{free}}(r, \alpha)$ and $\lim_{r\to\infty} \Pi^{\text{not free}}(r, \alpha) > \lim_{r\to\infty} \Pi^{\text{free}}(r, \alpha)$, i.e., free contracts can only be optimal if r is not too large and not too small.

Remark 2. With homogeneous cost $\bar{\xi} > 0$, $\underline{r}$, r^{free}, $\overline{r}$ and $r^{\text{not free}}$ correspond to $\frac{CF^*}{1-\alpha}$, $\bar{\xi}$, $\frac{\bar{\xi}-CF^*}{\alpha}$, and $\frac{\bar{\xi}}{\alpha}$, respectively. In Appendix B.4, we formalize this correspondence by considering a limit of models with heterogeneous costs converging to the one with the homogeneous cost.

Table 2 summarizes the results of Lemma 5 and Theorem 2 for the case when $\frac{G(\xi)}{g(\xi)}$ is convex.

B.2 Comparative Statics

Deriving precise comparative statics in the heterogeneous setup is daunting. While it is straightforward to show that $\Pi^{\text{not free}}(r, \alpha)$ and $\Pi^{\text{free}}(r, \alpha)$ are increasing in the size of externalities (r) and the fraction of the high types (α), it is hard to pin down how the comparison between these two values are affected as we change parameters (r and α). Nevertheless, using the partial characterization of the optimal contracts we can make comparative statics to understand robustness and changes of our results with the introduction of heterogeneity of WoM costs.

Proposition 7 (Market Structure and Free Contracts)**.** *The following claims hold in the model with heterogeneous costs for any fixed* $r \in [0, \infty)$*.* $\lim_{\alpha \to 0} \Pi^{not\ free}(r, \alpha) > \lim_{\alpha \to 0} \Pi^{free}(r, \alpha)$ *and* $\lim_{\alpha \to 1} \Pi^{not\ free}(r, \alpha) > \lim_{\alpha \to 1} \Pi^{free}(r, \alpha)$.[28]

The intuition for Proposition 7 is as follows. The only reason to offer free contracts is to boost up the expected externalities by $(1 - \alpha)r$, and such boosting is not significant if α is high, hence offering free contracts is suboptimal in those cases. With homogeneous costs, we showed in Section 3 that free contracts are optimal only when α is small. Similarly, with heterogeneous costs, a free contract cannot be optimal for high α. Moreover, if α is too small, $\Pi^{\text{free}}(r, \alpha) < 0$ holds because there are too few high types to compensate for the high cost of free contracts, and $\Pi^{\text{not free}}(r, \alpha) > 0$ holds because a strictly positive share of senders with very small WoM cost talk by part 1 of Theorem 2. This effect was not present with homogeneous costs, where the seller does not incentivize WoM at all, resulting in $\Pi^* = 0$.

The previous arguments imply that if there exists a set of parameters such that free contracts are optimal, then the choice of free versus non-free contracts is non-monotonic with respect to both r and α.

The comparative statics of the optimal reward scheme is more intricate with heterogeneous costs of WoM as the sender can fine-tune the number of senders that she wants to incentivize to engage in WoM.

Proposition 8 (Optimal Reward Scheme)**.** *Let* $r < r^{free}$*. Then, the following hold in the model with heterogeneous costs:*

(i) $R^{free}(r, \alpha)$ *is increasing in* α*.* $R^{not\ free}(r, \alpha)$ *is increasing in* α *if and only if* $\alpha r \hat{G}'(\alpha r + R^{not\ free}(r, \alpha)) < \Pi^{classic}$*, where we define* $\hat{G}(\xi) \equiv \frac{G(\xi)}{g(\xi)}$ *for all* $\xi \in \mathbb{R}_+$*.*

(ii) $R^{free}(r, \alpha)$ *and* $R^{not\ free}(r, \alpha)$ *are decreasing in* r*.*

(iii) Referrals and free contracts are strategic substitutes, i.e. $R^{free}(r, \alpha) < R^{not\ free}(r, \alpha)$ *for all* $r \in (0, r^{not\ free})$ *and* $\alpha \in (0, 1)$*.*

[28] These limits exist because of the monotonicity in α.

Although part (ii) has the same prediction as in the case with homogeneous WoM costs, the prediction in part (i) is different. We first explain the comparative statics regarding $R^{\text{free}}(r, \alpha)$. Under homogeneous costs, every sender talks and every receiver buys anyway under the usage of free contracts, so α does not affect the optimal reward level. With heterogeneous costs, however, the firm needs to tradeoff the gain and loss of increasing the rewards. The gain is the additional receivers who hear from the senders who start talking due to the increase of the rewards. The loss is the additional payments. The gain is increasing in α, so the firm has more incentive to raise the rewards.

The relationship of the optimal reward and α conditional on no free contracts being offered is ambiguous because two forces are present. First, higher α means more benefit from the receivers, and this contributes to the incentive to raise the rewards. On the other hand, higher α means more expected externalities, so there is less need to bribe a given sender. This contributes to lowering the rewards. Naturally, the second effect dominates when senders are relatively homogeneous, and indeed the optimal reward is strictly decreasing when G is completely homogeneous as in the main analysis. To formalize this idea, define

$$HMG \equiv \sup_x \left(\frac{G}{g}\right)'(x)$$

which can be interpreted as a measure of homogeneity of costs. If HMG is large, it means that there is a small range of costs of WoM that are held by many senders and HMG goes to infinity in the limit as G converges to the completely homogeneous one in (4). An implication of the condition in part (i) of Proposition 8 is that there exists $\overline{HMG} > 0$ such that if $HMG < \overline{HMG}$, then $R^{\text{not free}}(r, \alpha)$ is increasing in α.

Recall that both free contracts and positive rewards are used if and only if $r \in [\frac{CF^*}{1-\alpha}, r^{\text{free}})$.

Proposition 9 (Market Structure and Using Both Rewards and Free Contracts)**.** *The following claims hold in the model with heterogeneous costs:*

1. $\frac{CF^*}{1-\alpha}$ *and* r^{free} *are strictly increasing in* α.

2. $\frac{CF^*}{1-\alpha}$ *is strictly increasing and* r^{free} *is strictly decreasing in* c*.*

As in the homogeneous-cost model, free contracts can only be optimal if the size of externalities r is larger than $\frac{CF^*}{1-\alpha}$. Since this number is increasing in α, free contracts are optimal for small r in niche markets with small α. Thus, free contracts and referral rewards should be jointly used in niche markets (small α) if externalities are rather small, while they should be used in mass (larger α) markets if externalities are comparably larger.

With homogeneous costs, all receivers use the product under free contracts. Thus, what corresponds to r^{free} (which is $\bar{\xi}$) does not vary with α or c. With heterogenous costs, however, it varies with these parameters. This is because the increase in α or decrease in c contributes to an increase of the expected profit per receiver, which increases the firm's incentive to offer referral rewards.

B.3 Proofs

Proof. **(Lemma 5)** First, we show the existence of unique cutoffs r^{free} and $r^{\text{not free}}$. The first-order condition of $\Pi^{\text{free}}(r, \alpha)$ with respect to R is that (i) $R = 0$ or (ii) $R > 0$ and

$$g(r+R) \cdot \left[\pi((0, \underline{q}), (\tilde{p}_H^*, q_H^*)) - R - \frac{G(r+R)}{g(r+R)}\right] = 0.$$

Note that the expression in the bracket on the left-hand side is strictly decreasing given Assumption 4 and vary continuously from ∞ to $-\infty$ as R varies from $-\infty$ to ∞. Hence, the optimal reward is always unique in $\mathbb{R}$. Also, the same argument implies that there exists a unique r such that $\pi((0, \underline{q}), (\tilde{p}_H^*, q_H^*)) - \frac{G(r)}{g(r)} = 0$. Let this unique r be r^{free}. That is, the left-hand side of the first-order condition is nonpositive and thus $R^{\text{free}}(r, \alpha) = 0$ if and only if $r \geq r^{\text{free}}$.

Analogously, conditional on offering no free contracts ($q_L = 0$), the optimal reward is unique in $\mathbb{R}$ and there exists a unique r such that $\pi((0,0), (p_H^*, q_H^*)) - \frac{G(\alpha r)}{g(\alpha r)} = 0$. We denote this r by $r^{\text{not free}}$. As before, we have that $R^{\text{not free}}(r, \alpha) = 0$ if and only if $r \geq r^{\text{not free}}$.

Finally, we show that $r^{\text{free}} < r^{\text{not free}}$. To see this, note that Assumption 4 implies $\frac{G(\alpha r)}{\alpha r} < \frac{G(r)}{r}$ for $r > 0$ and $\alpha \in (0,1)$. Together with $\pi((0,0),(p_H^*,q_H^*)) > \pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*))$, $r^{\text{free}} < \alpha r^{\text{not free}}$ follows by Assumption 4 and the definitions of r^{free} and $r^{\text{not free}}$. Since $\alpha < 1$, this implies $r^{\text{free}} < r^{\text{not free}}$. □

Proof. **(Theorem 2)**

1. By Assumption 3, $\pi((0,0),(p_H^*,q_H^*)) > 0$ holds. Also, since $g(\xi) > 0$ for all $\xi \in \mathbb{R}_+$, $G(\xi) > 0$ for all $\xi > 0$. Hence, for any $r \in [0,\infty)$ and $\alpha \in (0,1)$, $[\pi((0,0),(p_H^*,q_H^*)) - R] \cdot G(\alpha r + R) > 0$ holds if $R \in (0, \pi((0,0),(p_H^*,q_H^*)))$. Thus, $\Pi^{\text{not free}}(r,\alpha) > 0$.

2. Note that the use of both, free contracts and positive rewards, is optimal only if $r < r^{\text{free}}$. Also, $r < r^{\text{free}}$ implies that rewards are positive. Furthermore, in that case the maximization problems defining $\Pi^{\text{free}}(r,\alpha)$ and $\Pi^{\text{not free}}(r,\alpha)$ both have inner solutions, so the two maximization problems can be rewritten as:

$$\begin{aligned} \Pi^{\text{free}}(r,\alpha) &= \max\nolimits_{x\in\mathbb{R}}(A^{\text{free}} - x)\cdot G(x) \\ \Pi^{\text{not free}}(r,\alpha) &= \max\nolimits_{x\in\mathbb{R}}(A^{\text{not free}} - x)\cdot G(x) \end{aligned} \tag{14}$$

   where $A^{\text{free}} = \pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*)) + r$ and $A^{\text{not free}} = \pi((0,0),(p_H^*,q_H^*)) + \alpha r$. Thus, $\Pi^{\text{free}}(r,\alpha) \geq \Pi^{\text{not free}}(r,\alpha)$ if and only if

$$\pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*)) + r \geq \pi((0,0),(p_H^*,q_H^*)) + \alpha r.$$

   This is equivalent to $r \geq \frac{CF^*}{1-\alpha}$. Also, by part 1 of the current theorem, $\Pi^{\text{free}}(r,\alpha) \geq \Pi^{\text{not free}}(r,\alpha)$ implies $\Pi^{\text{free}}(r,\alpha) > 0$. Overall, there exists an optimal scheme such that both free contracts and positive rewards are used if and only if $r \in [\frac{CF^*}{1-\alpha}, r^{\text{free}})$.

3. Consider a variable

$$\frac{\Pi^{\text{free}}(r,\alpha)}{\Pi^{\text{not free}}(r,\alpha)}. \tag{15}$$

This variable is well-defined because the denominator is always strictly positive by part 1 of the current theorem.

Step 1: Note that for $r \geq r^{\text{not free}}$, Lemma 5 shows that the rewards are zero in any optimal scheme. Hence, $\Pi^{\text{free}} = \pi((0,\underline{q}),(\tilde{p}_H^*, q_H^*)) \cdot G(r)$ and $\Pi^{\text{not free}} = \pi((0,0),(p_H^*, q_H^*)) \cdot G(\alpha r)$ hold, and thus (15) is differentiable with respect to r. If $\frac{G}{g}$ is convex, then

$$\frac{\partial}{\partial r}\frac{\Pi^{\text{free}}(r,\alpha)}{\Pi^{\text{not free}}(r,\alpha)} = \frac{\partial}{\partial r}\frac{\pi((0,\underline{q}),(\tilde{p}_H^*, q_H^*)) \cdot G(r)}{\pi((0,0),(p_H^*, q_H^*)) \cdot G(\alpha r)} =$$
$$\frac{\left(\frac{G(\alpha r)}{g(\alpha r)} - \alpha \frac{G(r)}{g(r)}\right) \cdot \pi((0,\underline{q}),(\tilde{p}_H^*, q_H^*)) \cdot g(r) \cdot g(\alpha r)}{[\pi((0,0),(p_H^*, q_H^*))] \cdot G(\alpha r)^2} < 0.$$

Thus, when $r \geq r^{\text{not free}}$, either (i) free contracts are not optimal for any $r \in [r^{\text{not free}}, \infty)$, or (ii) there exists a $\overline{r}' \geq r^{\text{not free}}$ such that there exists an optimal scheme in which free contracts are offered for $r \in [r^{\text{not free}}, \overline{r}']$, and no free contracts are offered under any optimal scheme for $r > \overline{r}'$. It must be the case that $\overline{r}' < \infty$ because

$$\lim_{r\to\infty} \frac{\pi((0,\underline{q}),(\tilde{p}_H^*, q_H^*)) \cdot G(r)}{\pi((0,0),(p_H^*, q_H^*)) \cdot G(\alpha r)} = \frac{\pi((0,\underline{q}),(\tilde{p}_H^*, q_H^*))}{\pi((0,0),(p_H^*, q_H^*))} < 1.$$

We let $\bar{r} = \bar{r}'$ in case (ii).

Step 2: Next, we consider the following three different cases: $\frac{CF^*}{1-\alpha} < r^{\text{free}}$, $\frac{CF^*}{1-\alpha} \in [r^{\text{free}}, r^{\text{not free}}]$, and $\frac{CF^*}{1-\alpha} > r^{\text{not free}}$.

- Let $\frac{CF^*}{1-\alpha} < r^{\text{free}}$. Then, it follows from part 2 of the current theorem that no free contracts are offered for $r < \frac{CF^*}{1-\alpha}$ and free contracts are offered for $r \in [\frac{CF^*}{1-\alpha}, r^{\text{free}}]$. For $r \in [r^{\text{free}}, r^{\text{not free}}]$,

$$\frac{\partial}{\partial r}\frac{\Pi^{\text{free}}(r,\alpha)}{\Pi^{\text{not free}}(r,\alpha)} = \frac{\partial}{\partial r}\frac{[\pi((0,\underline{q}),(p_H^*, q_H^*))] \cdot G(r)}{\max_{R\in\mathbb{R}}[\pi((0,0),(p_H^*, q_H^*)) - R] \cdot G(\alpha r + R)} =$$
$$\frac{\left(\frac{G(\alpha r + R^{\text{not free}}(r,\alpha))}{g(\alpha r + R^{\text{not free}}(r,\alpha))} - \alpha \frac{G(r)}{g(r)}\right) \cdot \pi((0,\underline{q}),(p_H^*, q_H^*)) \cdot g(r) \cdot g(\alpha r + R^{\text{not free}}(r,\alpha))}{[\pi((0,0),(p_H^*, q_H^*)) - R^{\text{not free}}(r,\alpha)] \cdot G(\alpha r + R^{\text{not free}}(r,\alpha))^2}.$$

Note that $\Pi^{\text{not free}}(r,\alpha)$ is differentiable in r by the Envelope Theorem. Moreover, if $\frac{G(\alpha r+R)}{g(\alpha r+R)} - \alpha\frac{G(r)}{g(r)} < 0$, then $\alpha r + R < r$ because $\frac{G(\xi)}{g(\xi)}$ is increasing in ξ by Assumption 4. Moreover, $R^{\text{not free}}(r, \alpha)$ is differentiable in r by the implicit function theorem applied to the first-order condition of $\Pi^{\text{not free}}$ and, letting $\hat{G}(\xi) := \frac{G(\xi)}{g(\xi)}$ for all ξ,

$$\frac{\partial}{\partial r}R^{\text{not free}}(r,\alpha) = -\frac{\alpha\hat{G}'(\alpha r + R^{\text{not free}}(r,\alpha))}{1+\hat{G}'(\alpha r + R^{\text{not free}}(r,\alpha))} < 0.$$

Thus,

$$\begin{aligned}
\frac{\partial}{\partial r}\left(\frac{G(\alpha r + R^{\text{not free}}(r,\alpha))}{g(\alpha r + R^{\text{not free}}(r,\alpha))} - \alpha\frac{G(r)}{g(r)}\right) &= \\
\alpha\left(\hat{G}'(\alpha r + R^{\text{not free}}(r,\alpha)) - \hat{G}'(r)\right) + \hat{G}'(\alpha r + R^{\text{not free}}(r,\alpha))\frac{\partial}{\partial r}R^{\text{not free}}(r,\alpha) &= \\
\alpha\left(\hat{G}'(\alpha r + R^{\text{not free}}(r,\alpha)) - \hat{G}'(r) - \frac{\hat{G}'(\alpha r + R^{\text{not free}})^2}{1+\hat{G}'(\alpha r + R^{\text{not free}}(r,\alpha))}\right) &< 0.
\end{aligned}$$

Thus, if the derivative of (15) is negative at $r' \in [r^{\text{free}}, r^{\text{not free}}]$ then (15) is decreasing for all $r \in [r', r^{\text{not free}}]$. Together with Step 1, this implies the following. In case (i), there exists $\overline{r} \in [r^{\text{free}}, r^{\text{not free}})$ such that free contracts are offered in an optimal scheme if and only if $r \in [\frac{CF^*}{1-\alpha}, \overline{r}]$. In case (ii), the current analysis shows that it is optimal to offer free contracts for all $r \in [r^{\text{free}}, r^{\text{not free}}]$, so free contracts are offered if and only if $r \in [\frac{CF^*}{1-\alpha}, \overline{r}]$, where $\overline{r}$ is the variable that we defined in Step 1.

- Let $\frac{CF^*}{1-\alpha} \in \left[r^{\text{free}}, r^{\text{not free}}\right]$. In that case, offering free contracts is not optimal for any $r < r^{\text{not free}}$. Then, either free contracts are not optimal for any r or by the same argument as above, if free contracts are not used in an optimal scheme for $r = r'$ then they are not used in any optimal scheme for any $r > r'$. This proves the desired claim for this case.
- If $\frac{CF^*}{1-\alpha} > r^{\text{not free}}$, then offering free contracts is not optimal for any $r < r^{\text{free}}$. For

$r \in \left[r^{\text{free}}, r^{\text{not free}}\right]$ free contracts are also not optimal because

$$1 > \frac{\max_{R\in\mathbb{R}}[\pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*)) - R]\cdot G(r+R)}{\max_{R\in\mathbb{R}}[\pi((0,0),(p_H^*,q_H^*)) - R]\cdot G(\alpha r+R)} \geq \frac{[\pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*))]\cdot G(r)}{\max_{R\in\mathbb{R}}[\pi((0,0),(p_H^*,q_H^*)) - R]\cdot G(\alpha r+R)}.$$

The first inequality follows from the proof of part 2 of the current theorem. For $r \geq r^{\text{not free}}$, offering free contracts is never optimal by Step 1.

This concludes the proof.

□

Proof. **(Proposition 7)** First, note that we can write the limiting profits as

$$\begin{aligned}
\lim_{\alpha\to 1} \Pi^{\text{free}}(r,\alpha) &= \max_{x\geq r}(\tilde{p}_H^* - cq_H^* + r - x)G(x) < \\
\lim_{\alpha\to 1} \Pi^{\text{not free}}(r,\alpha) &= \max_{x\geq r}(p_H^* - cq_H^* + r - x)G(x).
\end{aligned}$$

It follows immediately from part 1 of Theorem 2 that there exist $\alpha' > 0$ and $\epsilon > 0$ such that $\Pi^{\text{free}}(r,\alpha) + \epsilon < \Pi^{\text{not free}}(r,\alpha)$ for any $\alpha \in (0,\alpha')$, hence the limit result as $\alpha \to 0$ holds. □

Proof. **(Proposition 8)** Applying the implicit function theorem to the first-order conditions of Π^{free} and $\Pi^{\text{not free}}$ gives us:

(i) $\frac{R^{\text{free}}(r,\alpha)}{\partial\alpha} = -\frac{p_H^* - q_H^* c - v_H(\underline{q}) + c\underline{q}}{-1-\hat{G}'(r+R)} > 0$ and

$$\frac{\partial R^{\text{not free}}(r,\alpha)}{\partial\alpha} = -\frac{p_H^* - q_H^* c - r\hat{G}'(\alpha r+R)}{-1-\hat{G}'(r+R)}$$

which is strictly greater than zero if and only if $r\hat{G}'(\alpha r + R^{\text{not free}}(r,\alpha)) < p_H^* - q_H^* c$, or $\alpha r\hat{G}'(\alpha r + R^{\text{not free}}(r,\alpha)) < \Pi^{\text{classic}}$.

(ii) $\frac{R^{\text{free}}(r,\alpha)}{\partial r} = -\frac{-\hat{G}'(R+r)}{-1-\hat{G}'(R+r)} < 0$ and $\frac{R^{\text{not free}}(r,\alpha)}{\partial r} = -\frac{-\alpha\hat{G}'(R+\alpha r)}{-1-\hat{G}'(R+\alpha r)} < 0$ because $\hat{G}'(x) > 0$ for all $x > 0$, so $-1 - \hat{G}'(x) < 0$.

(iii) First, note that for $r > r^{\text{free}}$, referral rewards are always zero when free contracts are offered, i.e., the statement is trivially true. If $r \leq r^{\text{free}}$, then the optimal reward with free contracts $R^{\text{free}}(r,\alpha)$ satisfies the first-order condition:

$$R^{\text{free}}(r,\alpha) = \pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*)) - \frac{G(r+R^{\text{free}}(r,\alpha))}{g(r+R^{\text{free}}(r,\alpha))}. \tag{16}$$

By the first-order condition for the maximization problem for the case with no free contracts with respect to the reward, the solution $R^{\text{not free}}(r,\alpha)$ must satisfy:

$$g(\alpha r + R^{\text{not free}}(r,\alpha)) \cdot \left(\pi((0,0),(p_H^*,q_H^*)) - R^{\text{not free}}(r,\alpha) - \frac{G(\alpha r + R^{\text{not free}}(r,\alpha))}{g(\alpha r + R^{\text{not free}}(r,\alpha))}\right) = 0.$$

Since $g(\cdot) > 0$, this implies that

$$\pi((0,0),(p_H^*,q_H^*)) - R^{\text{not free}}(r,\alpha) - \frac{G(\alpha r + R^{\text{not free}}(r,\alpha))}{g(\alpha r + R^{\text{not free}}(r,\alpha))} = 0. \tag{17}$$

Now, substitute $R^{\text{not free}}(r,\alpha)$ by the expression for $R^{\text{free}}(r,\alpha)$ given by (16) on the left hand side of (17), to obtain:

$$\pi((0,0),(p_H^*,q_H^*)) - \pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*)) + \frac{G(r+R^{\text{free}}(r,\alpha))}{g(r+R^{\text{free}}(r,\alpha))} - \frac{G(\alpha r+R^{\text{free}}(r,\alpha))}{g(\alpha r+R^{\text{free}}(r,\alpha))}.$$

This is strictly positive by log-concavity of G (Assumption 4) and because $\pi((0,0),(p_H^*,q_H^*)) > \pi((0,\underline{q}),(\tilde{p}_H^*,q_H^*))$. Noting that the left hand side of (17) is strictly decreasing in referral rewards, the optimal reward without free contracts $R^{\text{not free}}(r,\alpha)$ is strictly greater than $R^{\text{free}}(r,\alpha)$. □

Proof. **(Proposition 9)** The comparative statics with respect to $\frac{CF^*}{1-\alpha}$ are straightforward from the formula of CF^*. The ones for r^{free} follow from the first-order condition with respect to rewards that appears in the proof of Lemma 5 and Assumption 4. □

B.4 Homogeneous Costs as the Limit of Heterogeneous Costs

Consider a sequence $\{G^n\}_1^\infty$ that converges pointwise to (4) such that for each n, G^n is twice differentiable with $(G^n)'(\xi) = g^n(\xi) > 0$ for all ξ, and Assumption 4 holds. Let the set of all such sequences be $\mathcal{G}$. The set $\mathcal{G}$ is nonempty. For example, consider $\{G^n\}_1^\infty$ such that or each $n \in \mathbb{N}$, G^n is a normal distribution with mean $\bar{\xi} \geq 0$ and variance $\frac{1}{n}$ truncated at $\xi = 0$. By inspection one can check that $\{G^n\}_1^\infty \in \mathcal{G}$. For any given G^n, we can define $\underline{r}^n$, $r^{\text{free},n}$, $\bar{r}^n$, and $r^{\text{not free},n}$. Then, the following statement can be shown: For any $\{G^n\}_1^\infty \in \mathcal{G}$,

$$\lim_{n\to\infty} \underline{r}^n = \frac{CF^*}{1-\alpha}, \quad \lim_{n\to\infty} r^{\text{free},n} = \bar{\xi}, \quad \lim_{n\to\infty} \bar{r}^n = \frac{\bar{\xi} - CF^*}{\alpha}, \quad \text{and} \quad \lim_{n\to\infty} r^{\text{not free},n} = \frac{\bar{\xi}}{\alpha}.$$