PRELIMINARY AND INCOMPLETE

PLEASE DO NOT CITE WITHOUT PERMISSION FROM THE AUTHORS

Consumer Tracking and Efficient Matching in Online Advertising Markets

Susan C. Athey, Emilio Calvano and Joshua S. Gans

December 31, 2012

One of the advances in our understanding of two-sided markets or platforms is the notion of a *competitive bottleneck*. This arises in the context of competing platforms when a group on one side of the market always multi-homes; that is, they pay to access each platform. More strongly, they pay to access and use any one platform independently of what they are doing on other platforms. This has a significant impact on the nature of competition between platforms. If one side always multi-homes then it is, in some sense, captured by each platform. The platform can act like a monopolist with respect to those customers. However, this does not eliminate competition as, by the very nature of two-sided markets, the level of activity on the other side of the market impacts on the quality of the product served to the other. In this situation, platforms may compete for increased use on one side of the market in order to increase the supply and quality of the product they can provide to captured customers.

This model is pervasive in two-sided markets involving advertisers. The canonical model of media economics (Anderson and Coate, 2005) has, as its baseline, a model where each advertiser wants to communicate to each broadcast viewer. Consequently, competing broadcasters can charge advertisers a monopoly price for access to any given viewer. The broadcasters then compete over the number of viewers on their channel. An important implication of this is that if broadcasters were to merge, there would be no change in prices or welfare in the advertising-side of the market although there would be a reduction in competition for viewers. This type of analysis has played a significant role in anti-trust analysis involving advertising markets and mergers.

For this reason, it is important to understand the behavior of advertisers with respect to single or multi-homing or something in between. The baseline view is that advertisers want to send messages to consumers and, indeed, place a value on impressing each and every one of them. Thus, if a platform happens to have attracted a consumer, then an advertiser must deal with the platform to send a message to that consumer. Thus, multi-homing behavior is a direct implication of a desire to send a message to all consumers. Of course, this assumption might be relaxed if, for instance, advertisers are not interested in such complete coverage. For instance, an advertiser may have a limited quantity of products to sell or may have another limit on their marketing budget. This will constrain their behavior in sending messages to all consumers.

Another factor is that the platforms may not capture a consumer entirely. As modelled by Ambrus, Calvano and Reisinger (2012), some consumers may consume both platforms weakening platforms claim to capture those consumers and leading to direct competition for them. Competition for captured consumers would not arise. Or alternatively, as modelled by Athey, Calvano and Gans (2012), some share of the consumer market may allocate attention across platforms while others may concentrate it on platforms but it may not be possible ex ante to identify particular consumers by their single and multi-homing behavior. In this case, all consumers are potentially contestable and this will impact on advertiser behavior in terms of single or multi-homing. In each case, a merger between platforms may reduce the competition for advertisers.

Here we present an alternative driver of advertiser behavior that we believe is of relevance for online advertising markets. The markets we have in mind are newly emerging ad networks that offer display advertising for many web-pages and outlets. Unlike traditional media (such as newspapers), at a base case, we can assume that consumers location on a platform is not known ex ante; that is, there are no captured consumers as ad networks cannot be assumed to drive a consumer's choice of content. In this sense, from the perspective of an ad network, all consumers are multi-homers.

By abstracting away from the standard, assumed behavior of consumers, we are able to identify a new driver of advertiser behavior: the technology by which messages are communicated to consumers. Put simply, the nature of the communicating technology impacts on the efficiency of choices advertisers face in allocating messages across platforms. In the process, we demonstrate that the type of technology available or adopted will impact on conclusions regarding advertiser multi-homing.

We model communication as a stochastic process in which a sender (for example an advertiser) transmits a number of (costly) messages (for example ads) to a set of receivers (the consumers) through two different channels (the outlets). The process determines who (and hence how many) get at least one message (and hence are informed) if the sender transmits a certain number of messages through given outlet. Such stochastic process is what we call a "communication technology." Different outlets supply one such technology. It can be efficient, in which case each additional message hits an uniformed receiver. Or it can be dumb, in which case each additional message hits a random receiver, possibly an already informed one. Or it can be anything in between. Communication is costly in the sense that messages are unit-priced.

The paper proceeds as follows. In Section 2, we present our model framework embedded within a simple theory of communication. In Section 3, we consider our baseline case with identical outlets both of which share the same technology and the same set of receivers. This allows us to characterize the impact of tracking technology on advertiser multi-homer behavior. In Section 4, we relax the latter assumption and consider how tracking technology interacts with the share of overlapping receivers outlets have. In Section 5, we consider asymmetric outlets – specifically, a case where one outlet's consumers are a subset of the other's. This allows us to consider trade-offs between reach and readership and the impact of this on advertiser behavior. We then turn to consider the welfare implications of advertiser behavior; particularly, on receivers. This is done by recognizing that wasted messages are likely to cost receivers attention with no other potential benefits. We provide some speculative and illustrative discussion on these welfare effects. A final section concludes.

1 A simple theory of communication

Consider a set (measure one) of senders, a set (measure one) of receivers and two outlets. Outlets are identical but for their set of receivers denoted respectively R_1 and R_2 and assumed to be of equal measure. The measure of $R_1 \cup R_2$ is 1. Let D^s and D_i^l denote the measure of $R_1 \cap R_2$ and $R_i \setminus (R_1 \cap R_2)$ respectively (s stands for switcher and l for loyal). With an abuse of notation we denote with $|\cdot|$ the operator that maps a set to its measure. Senders wish to inform receivers. The unit price of a message on either outlet is $p_i > 0$. If p_i is equal to p_j then we use p to denote the common prce. Let v denote the expected value of informing a receiver. Type v's payoff when choosing to purchase (n_1, n_2) messages is v times the expected fraction of the population informed minus communication costs $p_1 * n_1 + p_2 * n_2$.

A communication (or tracking) technology is a function $\psi_i : \mathbb{R}_+ \to [0, 1]$ that maps the number of messages *per receiver* sent through a given outlet i to the probability of informing a given receiver at least once. If there are D receivers in total then $\psi_i(n/D)$ is also the expected fraction of the receivers reached by at least one message out of the n messages sent through outlet i. We assume ψ_i strictly increasing on $[0, 1)$ (if $\psi_i < 1$, constantly equal to 1 otherwise) and concave. Finally, we assume that $\psi_i(n') - \psi_i(n'') \leq n' - n''$

for all $n', n'' \geq 0$. That is, one message can hit at most one receiver. This implies that ψ_i cannot increase at a rate greater than one and that $\psi_i(0) = 0$.

One way to think about these assumptions is as reflecting a minimum level of intelligence or efficiency of the communication process. An extra message should have some chance of informing. Concavity captures intelligence in a subtler way. Any strategy that can help identify uninformed receivers is exploited as soon as possible. These assumptions also rule out learning by doing.

The above implicitly restricts attention to communication processes with homogeneous receivers in the sense that no receiver is ex-ante more likely to be informed than the others. Of course, that need not be true ex-interim, since each additional message is a random trial whose realization can depend on the realization of the previous trials. For example, an uniformed receiver might be more likely to be reached by a given message as the number of already informed receivers increases.

In what follows we introduce advertisers' payoffs and characterize their choices as a function of the technology employed and the distribution of receivers accross outlets. Rather then deploying the more general model upfront, we build our theory step by step. We first study the simplest degenerate case in which both outlets share the same set of receivers. We then move on to the more complex cases in which receivers partly overlap.

2 The advertiser's (symmetric) dilemma

2.1 Identical Outlets

We consider a situation where no one consumer can be identified exclusively with a particular outlet. As mentioned earlier, this might arise because each consumer may consume content on outlets utilizing any ad network. Thus, we suppose that $R_1 = R_2$ with $|R_i| = 1$ and $\psi_i(n) = \psi_j(n) =: \psi(n)$. (figure 1)

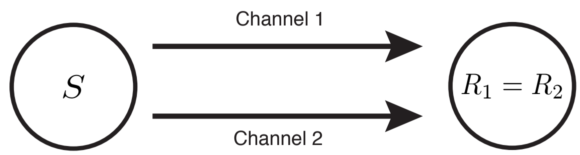

Fig. 1: Imperfect communication with full overlap.

A first observation is that the fraction of the population informed through (n_1, n_2) messages is equal to one minus the probability that a given receiver is not informed through either outlet. Under independence, that is $1-(1-\psi(n_1))(1-\psi(n_2))$. Type v's choice is determined by:

$$\max_{n_1,n_2\geq 0} v(1-(1-\psi(n_1))(1-\psi(n_2))) - p(n_1+n_2). \tag{1}$$

Let (n_1^*, n_2^*) denote a solution to the above problem and $n^* := n_1^* + n_2^*$ be the total number of impressions purchased across all outlets. We say that a non-trivial ($n^* > 0$) solution point is characterized by **maximum diversification** if $n_1^* = n_2^*$ and **maximum concentration** if either n_1^* or n_2^* is equal to n^*. Consider the family of optimization problems parametrized by v and let $n_1(v)$, $n_2(v)$ and $n(v) := n_1(v)+n_2(v)$ denote the mapping form v to the solution.

Proposition 1. *For all values of $v > 0$ there is a unique solution $n^* \geq 0$. In addition $n(v)$ is monotone increasing in v.*

Proposition 2. *a) If $1-\psi$ is log-concave then there is maximum concentration at all solution points.*
b) If $1-\psi$ is log-linear then all elements of $\{n_1, n_2 \geq 0 : n_1 + n_2 = n^*\}$ are solutions.
c) If $1-\psi$ is log-convex then there is maximum diversification.

Proposition 1 follows from the increasing differences property of the objective function. Higher types face a higher opportunity cost of not informing receivers. This implies that higher types will transmit more messages in total (if any).[1] To sketch the argument used to prove Proposition 2 it is useful to study how impressions on different outlets substitute for one another under technology ψ. That is, consider the following problem:

$$\max_{n_1,n_2\geq 0}(1-(1-\psi(n_1))(1-\psi(n_2))) \qquad \text{s.t.} \quad n_1+n_2=n.$$

A crucial observation is that log-concavity, log-convexity and log-linearity pin down the sign of the rate of change of the marginal rate of substitution of impressions on different outlets. In particular, if $1-\psi$ is log-concave, then the marginal rate of substitution is increasing (Figure 1, left). In terms of the "optimal allocation" problem presented above log-concavity calls for concentrating all impressions on one outlet. A simple illustration of a log-concave communication technology is perfect internal tracking: $\psi(n)=\inf\{n,1\}$.[2] Under log-linearity, impressions on different outlets are perfect substitutes when it comes to impress the shared receivers. Figure 2, center, illustrates the level curves for $\psi(n)=1-e^{-n}$ (the so-called Butters (1977) "no tracking" technology). Finally, log-convexity implies strictly convex upper-contour sets. (Figure 2, right). The solution to this problem is characterized by a usual tangency condition between a level curve and the budget set $\{n_1,n_2\geq 0 : n_1+n_2=n\}$.

To build intuition, notice that there are two sources of missed and wasted messages in this framework. First, additional messages transmitted through the same outlet can reach an already informed receiver and hence get wasted. This is what we call **within outlet waste**. Precisely, there is no within outlet waste if $\psi(n)=\inf\{n,1\}$ for all $n\geq 0$. There is within outlet waste otherwise. Second, because of independence, two messages sent on different outlets can reach the very *same* receiver even when there is no within outlet waste. This is what we call **across outlet waste**. Precisely, there is

[1]In fact if some type v purchases a positive quantity, then all higher types will purchase a *strictly* higher quantity.

[2]Log(1-x) is concave.

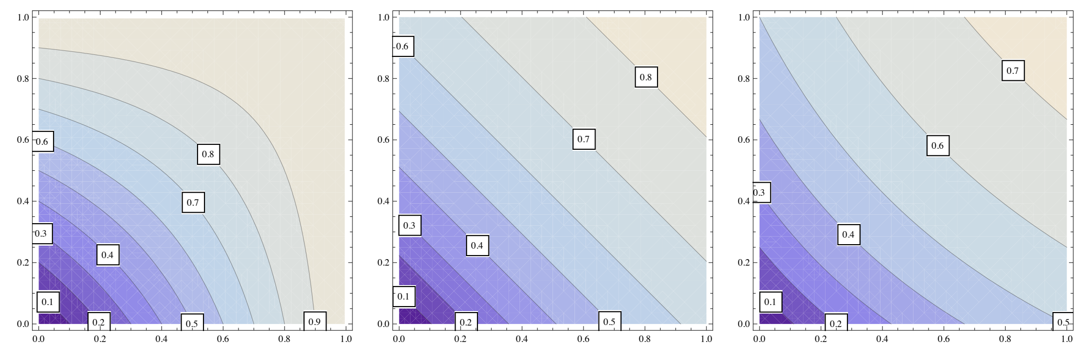

Figure 2: Level curves of $1-(1-\psi(n_1))(1-\psi(n_2)$. From left to right $1-\psi$ is log-concave, log-linear and log-convex respectively.

no across outlet waste whenever $1-(1-\psi(n_1))(1-\psi(n_2)) = \inf\{n_1+n_2, 1\}$. There is across outlet waste otherwise. Notice that the assumption that one message hits at most one receiver implies that there is always across outlet waste even if $\psi(n) = n$ for all $n \leq 1$.

The curvature of the logarithm of $1-\psi$ captures the relationship between within outlet waste and across outlet waste. This point can be illustrated formally through the equivalence between log linearity and log additivity in this particular context.

Lemma 1. *$(1-\psi)$ is log-linear if and only if $(1-\psi)$ is log-additive, that is: $\log(1-\psi(n_1+n_2)) = \log(1-\psi(n_1)) + \log(1-\psi(n_2))$ which is equivalent to:*

$$\psi(n_1+n_2) = 1-(1-\psi(n_1))(1-\psi(n_2)) \text{ for all } n_1, n_2 \geq 0 \tag{2}$$

Proof. By lemma 4 in appendix we have that $(1-\psi(n_1))(1-\psi(n_2))$ is constant over the set $\{n_1, n_2 \geq 0 : n_1 + n_2 = n\}$ for all $n \geq 0$. This coupled with the assumption that $\psi(0) = 0$ implies that for all n we have

$(1-\psi(n))(1-\psi(0)) = k = (1-\psi(n_1))(1-\psi(n-n_1))$ which is equivalent to (2). $\square$

Equation (2) states the equivalence between the two sources of waste. Adapting the same logic we can characterize log-concavity and log-convexity. $1-\psi$ is log concave (log-convex) if and only if $\psi(n_1+n_2)$ is strictly greater (respectively strictly lower) than $1-(1-\psi(n_1))(1-\psi(n_2))$ for all n_1, n_2. Using this characterization we can restate Proposition 2 as follows:

Proposition 3. *a) If within outlet waste is lower than across outlet waste then there is maximum concentration at all solution points.*
b) If within outlet waste is equal to across outlet waste then all feasible allocations of n^* are solutions.
c) If within outlet waste is greater than across outlet waste then there is maximum diversification.

This gives us a new understanding of a driver in advertiser multi-homing behavior. For a given advertiser, v, and message price, p, the nature of the communication technology will drive whether that advertiser will want to concentrate their advertising on all outlets, be indifferent or alternatively diversify maximally. If they choose to concentrate their advertising then, if impressions sold for different prices on each outlet, advertisers will choose to purchase impressions on the lowest price outlet. Thus, there will be direct competition for advertising business and this price competition will only be mitigated by any limitations (as yet unmodelled) on the capacity of an outlet to deliver ads.

By contrast, under some communication technologies where within outlet waste is greater than across outlet waste, advertisers will choose to diversify their impressions across outlets. In this situation, it remains the case that a price differential can cause advertisers to place more impressions on the low price outlet but the elasticity of the response will be mitigated by the lower efficiency of the marginal impression to that outlet. When might within outlet waste be greater than across outlet waste? This might arise if there is

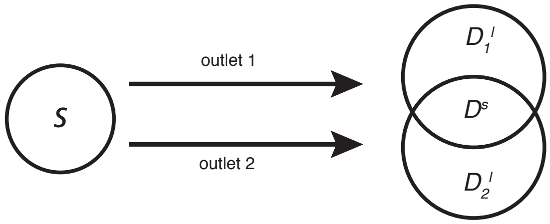

Figure 3: Imperfect communication with partial overlap.

no internal tracking capability but a typical consumer chooses to consume one outlet in the morning and the other in the afternoon; that is, there is content diversification that aggregates to an attractive bundle for consumers compared with the consumer choosing all the content of a particular outlet. This highlights the fact that part of the communciation technology could embed self-selection by consumers to tailored content.

2.2 Symmetric outlets with overlapping receivers

We now relax the assumption $R_1 = R_2$ allowing outlets to have their own captured set of receivers (see figure 3). Here receivers belong to two sets (or types): loyal of 1,2 and switchers. We retain the assumption that receivers are homogeneous in the sense that they only differ with regard to the outlet they are connected to. With partial overlap, capturing this simple idea requires an extra ingredient. Switchers get messages from two different sources. In order to avoid switchers being more likely to be informed because of this, we assume that any given switcher is twice less likely to receive a message sent through a given outlet.[3] Specifically, suppose a total of n_1 messages are transmitted through outlet 1. If loyals are twice as likely than

[3]For example, consider the following application. All consumers are endowed with two units of time. Some consumers choose to spend both units on outlet 1. Some consumers allocate both units to outlet two. Some consumers spend one unit of time on each outlet. If messages are sent at random times then a loyal is twice more likely than a switcher to be the recipient of a given message.

switchers to receive a message then the expected fraction of n_1 that goes to them is $\eta_1^l := 2D_1^l/(2D_1^l + D^s)$. (One way to think about this is as if each loyal receiver of outlet i were endowed with two mailboxes, each switcher receiver with one mailbox and messages on outlet i were randomly allocated to each mailbox in a uniform way; in the sense that no *mailbox* is ex-ante more or less likely to get a given message). Since by definition the argument of the ψ function is the average (per receiver) number of messages sent, then the corresponding fraction of D_1^l informed when sending a total of n_1 messages through channel i is equal to

$$\psi\left(\frac{\eta_1^l n_1}{D_1^l}\right) = \psi\left(\frac{2n_1}{2D_1^l + D^s}\right). \tag{3}$$

The above simplifies to $\psi(2n_1)$ when outlets are symmetric: $D_1^l = D_2^l =: D^l$. Iterating the same reasoning we find that switchers receive a total number of $n_1(1 - \eta_1^l) = n_1 D^s/(2D_1^l + D^s)$ from outlet one. Under symmetry, $2D_1^l + D^s = 1$ so the total number is $n_1 D^s$. That is n_1 messages per switcher. Symmetrically, if n_2 messages are sent through outlet 2 then n_2 per switcher reaches these latter. Since messages sent through different outlets are statistically independent, the expected number of switchers reached when sending (n_1, n_2) messages is $1 - (1 - \psi(n_1))(1 - \psi(n_2))$.

The sender's dilemma (normalizing by v the price p) is:

$$\max_{n_1,n_2\geq 0} 2D^l(\psi(2n_1)+\psi(2n_2))+D^s(1-(\psi(n_1))(1-\psi(n_2)))-(n_1+n_2)p/v \tag{4}$$

Note that if $D^s = 1$ (so that $D^l = 0$) we are back to the previous section's case. It is useful to decompose this problem in two sub-problems. First, given a total number of impressions n to allocate across outlets, which of all possible allocations maximizes reach? Second, given the solution to the above problem (denoted with star decorations), how many impressions should an advertiser of type v purchase if the price is the same for both

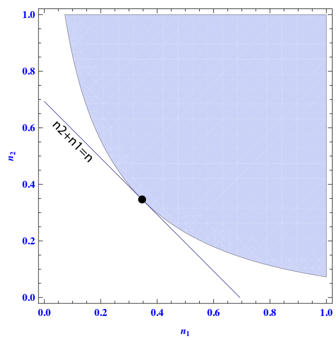

Figure 4: Optimal allocation for $D^s = 0$

outlets? Consider the first stage problem.

$$\max_{n_1,n_2 \geq 0} 2D^l(\psi(2n_1)+\psi(2n_2))+D^s(1-(1-\psi(n_1))(1-\psi(n_2))) \qquad \text{s.t. } n_1+n_2 = n. \tag{5}$$

Proposition 4. *If $1-\psi$ is log-concave there exists a threshold $0 < \tilde{D}^s < 1$ such that the solution to the optimal allocation problem entails maximum concentration if $D^s > \tilde{D}^s$ and full diversification if $D^s < \tilde{D}^s$. If $1-\psi$ is log-linear, for all $D^s < 1$ the solution to the optimal allocation problem entails full diversification.*

Consider firstly the case $1-\psi$ log-concave. If D^s were equal to zero then decreasing marginal returns on loyals would imply an even allocation of impressions across outlets: $n_1^* = n_2^* = n/2$. Indeed the level curves of $\psi(2n_1) + \psi(2n_2)$ are always convex towards the origin (strictly convex if $\psi'' < 0$) and the solution can be easily visualized as the tangency point at $n_1 = n_2$ (figure 4).

On the contrary if D^s is equal to one then we know that the optimal allocation calls for maximum concentration from the previous section. The threshold D^s is such that within outlet waste equals accross outlet waste so that the level curves of the objective function (5) have constant slope. In this case, the optimal allocation problem has infinite solutions.

Finally, given the optimal allocation $(n_1^*(n), n_2^*(n))$ the optimal total number of messages purchased is found by solving:

$$\max_{n\geq 0} D^l\psi(2n_1^*(n))+D^l\psi(2n_2^*(n))+D^s(1-(1-\psi(n_1^*(n)))(1-\psi(n_2^*(n))))-np/v \tag{6}$$

How does the choice of the sender change with D^s?

Proposition 5. *Suppose $\psi : \mathbb{R}_{++} \to [0,1]$ is continuous, a.e. differentiable, increasing, concave, bounded above by one and satisfies the following boundary conditions: $\psi(0) = 0$, $\lim_{n\to\infty} \psi(n) = 1$. If $(1-\psi)$ is log-linear then advertisers' choices do not depend on consumer switching.*

Thus, when $(1-\psi)$ is log-concave, advertisers' multi-homing choices interact with the share of consumers who are switchers. If there are no switchers, advertisers will choose maximum diversification because there is technically no across outlet waste as this depends on there being switchers. As the number of switchers rises, any across outlet waste will rise and weaken advertiser incentives to multi-home. Interestingly, this does not occur when $(1-\psi)$ is log-linear as within and across outlet waste are the same regardless of the number of switchers. If $(1-\psi)$ is log-concave we conjecture that there exists a threshold $\tilde{v}$ such that advertisers' demand for impressions decreases with D^s in $(p, \tilde{v})$ and increases with D^s in $(\tilde{v}, \infty)$. We have no yet demonstrated this conjecture for the general case here but it does emerge in the specific model in Athey, Calvano and Gans (2012).

3 The advertiser's asymmetric dilemma

Suppose $R_2 \subset R_1$ (Figure 5). That is, one set of receivers fully contains the other. In line with the above notation let $D^s = |R_2|$, $D_1^l = |R_1 \setminus R_2|$ and

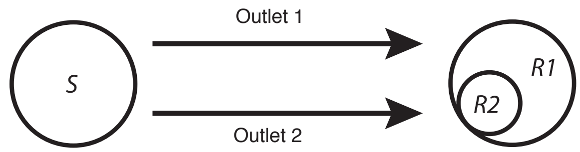

Fig. 5: Asymmetric outlets with $D_2^l = 0$

$D_2^l = 0$. Assume once more $p_1 = p_2 = p$. Then the advertiser's dilemma is (dividing everything by v):

$$\max_{n_1,n_2} D_1^l \psi\left(\frac{2D_1^l}{2D_1^l + D^s}\frac{n_1}{D_1^l}\right) + D^s\left(1 - \left(1 - \psi\left(\frac{D^s}{2D_1^l + D^s}\frac{n_1}{D^s}\right)\right)\left(1 - \psi\left(\frac{n_2}{D^s}\right)\right)\right) - (n_1 + n_2)p/v. \tag{7}$$

The first term are the expected revenues on loyals of outlet 1. $2D_1^l/(2D_1^l + D^s)$ is the expected fraction of the total number of messages n_1 that lands on the loyals of 1. This quantity is then normalized by the amount of loyals D_1^l. That gives the average number of messages that land on loyals; i.e., on receivers in $R_1 \setminus R_2$. It follows that the expected reach among these is $D_1^l \psi(2n_1/(2D_1^l + D^s))$. The second term represents expected revenues on shared receivers. It can be easily derived resorting to the same logic. As will be clear later on, for our purpose it is convenient to operate a change of variables and rewrite the above objective in terms of the average number of impressions. So suppose the control variables are

$$\tilde{n}_1 := \frac{n_1}{D_1^l + D^s/2} \qquad \text{and} \qquad \tilde{n}_2 := \frac{n_2}{D^s/2}.$$

Then (7) can be rewritten as follows:

$$\max_{\tilde{n}_1,\tilde{n}_2} D_1^l \psi(\tilde{n}_1) + D^s\left(1 - \left(1 - \psi\left(\frac{\tilde{n}_1}{2}\right)\right)\left(1 - \psi\left(\frac{\tilde{n}_2}{2}\right)\right)\right) - \left(\tilde{n}_1(D_1^l + D^s/2) + \tilde{n}_2(D^s/2)\right) p/v. \tag{8}$$

Proposition 6. *If ψ is log-linear then $\tilde{n}_1^*(p/v) = \tilde{n}_2^*(p/v)$ and decreases with p/v.*

Note that despite the asymmetric model there is a symmetric characterization of the equilibrium advertising policy provided the focus is shifted to the average number of messages sent to each subgroup.

Next we look at the case where accross outlet waste is not equal to within outlet waste. To gain insight on the problem we assume the following functional form:

$$\psi(n) = 2e^n/(1+e^n) - 1, \tag{9}$$

and specific parameters $p = 1$, $v = 10$ and one outlet assumed to be twice as large as the other. That is $D_1^l = 1/2$ and $D^s = 1/2$. The correponding optimal choices are $\tilde{n}_1 \approx 2.9$ while $\tilde{n}_2 \approx 3.4$. So the smaller outlet commands a higher demand in proportion to the number of its customers. The higher marginal returns on shared customers induce a higher average demand for messages to be delivered this subgroup.

More generally, in the appendix, we show that if $\psi(n) = 2e^n/(1+e^n) - 1$ and the outlets are asymmetric ($D_1^l \neq D_2^l$) then $\tilde{n}_1 \neq \tilde{n}_2$ whenever $p_1 = p_2$. We expect this property to hold for any log-concave ψ function but we have not proved it as yet.

The result here mirrors the *magnet content* discussion in Athey, Calvano and Gans (2012). Similar to this case, there one outlet only attracted switchers while the other outlet could have some exclusive consumers. The context we had in mind were sites like Facebook that had a large reach but did not necessarily capture the exclusive attention of consumers. We demonstrated that as the reach of the high reach outlet increased, it attracted more demand from single-homing advertisers who chose to advertise exclusively on the high reach outlet so as to eliminate across outlet waste. They could do this because the share of missed consumers would be relatively low. We saw this as indicative of the impact that imperfect tracking technologies can have on the type of content supplied by outlets; that is, whether the content emphasized reach versus attention per se.

4 Receivers' welfare and efficient communication

Thusfar, we have focussed on advertiser incentives to send messages and, in particular, their allocation across outlets. The primary cost to advertisers of wasted messages was the additional payments for those messages. However, there is also an additional cost that is not necessarily incurred by the advertisers; the attention of receivers. In this section we order the different advertising policies according to how wasteful these policies are from the point of view of the receivers. We sidestep the issue of comparing welfare changes accross senders and receivers by characterizing those policies that minimize the burden on receivers while keeping the senders' welfare fixed at some predetermined level.

Total waste (or duplicated impressions) equals the difference between the total number of messages sent and the total number of receivers hit by at least one message. If this latter quantity is denoted $r(n_1, n_2)$ then total waste is equal to:

$$n_1 + n_2 - r(n_1, n_2) \tag{10}$$

With this measure at hand we can ask if there is any wedge between what the sender chooses and what the recievers would like them to choose conditional on leaving them at least on the same utility level. That is, conditional on keeping the expected reach of the campaign equal or higher. We say that an advertising policy is k**-efficient** given prices (p_1, p_2) if (n_1^*, n_2^*) minimizes (10) subject to informing at least an expected fraction of the population $k \in (0, 1)$. Formally, given the following choice set:

$$\{(n_1, n_2) \in \mathbb{R}_+ : r(n_1, n_2) \geq k\} \tag{11}$$

an advertising policy (n_1^*, n_2^*) belonging to the above set is k-efficient if

(n_1^*, n_2^*) also minimizes (10) for all (n_1, n_2) belonging to the set; i.e., a self-interested and message-adverse receiver who therefore minimizes (10) subject to (11) would choose the same element of the set.

Proposition 7. *Suppose that the advertiser's optimal choice at prices (p_1, p_2) is (n_1^*, n_2^*) and the associated expected fraction of the population reached is k^*. Then (n_1^*, n_2^*) is k-efficient if and only if $p_1 = p_2$.*

Efficiency requires the marginal reach to be equalized across channels:

$$\frac{\partial r(n_1, n_2)}{\partial n_1} = \frac{\partial r(n_1, n_2)}{\partial n_2}. \tag{12}$$

Intuitively, if this condition were violated then it would be possible to reduce the total number of messages sent $n_1 + n_2$ while keeping the total reach of the campaign fixed by subsituting messages across outelts. The sender's optimal policy given unit prices p_1 and p_2 solves:

$$\frac{\partial r(n_1, n_2)}{\partial n_i} = \frac{p_i}{v} \qquad i = 1, 2. \tag{13}$$

That is, it equates the marginal reach to the unit price normalized by the value of informing a receiver. So the sender's optimal choice satisfies (13) if and only if $p_1 = p_2$. So it follows that any gap in unit prices will result in a wasteful allocation of messages from the perspective of receivers.

The next section supplies one (obvious) reason for why the prices, which we intepret as market clearing prices, need not be equal: log-concave tracking coupled with asymmetries across outlets.

5 Allocating messages through a simple market mechanism

We introduce here a simple market mechanism that allocates messages (and therefore the receivers' attention) to a population of identical senders of mass

1. In line with the above, we model receivers as passive and do not study the more complex problem in which the total amount of attention supplied to each outlet depends on the amount of messages (i.e. advertising) provided. Formally, each loyal receiver generates an inventory of β messages whereas each switcher generates an inventory of $\beta/2$ messages on each outlet. So the total supply of messages on outlet i is equal to $\beta(D_i^l + D^s/2)$. If messages are unit prices then the market clearing prices, denoted p_1^* and p_2^* solve:

$$\begin{aligned} n_1(p_1, p_2) &= \beta(D_1^l + D^s/2) \\ n_2(p_1, p_2) &= \beta(D_2^l + D^s/2) \end{aligned} \tag{14}$$

Or, alternatively, in average terms:

$$\begin{aligned} \tilde{n}_1(p_1, p_2) &= \beta \\ \tilde{n}_2(p_1, p_2) &= \beta. \end{aligned} \tag{15}$$

Let $n_1^* := n_1(p_1^*, p_2^*)$ and $n_2^* := n_2(p_1^*, p_2^*)$ denote a competitive allocation, that is the senders' demand evaluated at the market clearing prices.

A corollary of Proposition 6 is that if the technology is log-linear then the corresponding market clearing prices must be equal. By Proposition 7 then the competitive allocation must be k-efficient. However the same need not hold if ψ is log-concave. For instance, given the log-concave $1 - \psi$ function (9) we have that $\tilde{n}_1$ never equals $\tilde{n}_2$ when $p_1 = p_2$. Because market clearing requires $\tilde{n}_1 = \tilde{n}_2$, the *combination* of log-concave tracking technlogies and asymmetries accross outlets can potentially induce a wedge in the equilibrium prices and, therefore, an ineffcient allocation of the outlets' inventory.

Note that the market clearing prices reflect only (relative) scarcity considerations. Because the receivers' attention is an unpriced resource, there is no market mechanism at work to restore efficiency. This is despite there being a simple coincidence of interest between senders and receivers: the lower the waste the higher the amount of information that can be pushed to receivers.

6 Conclusions

To be done.

7 References

Ambrus, A., E. Calvano and M. Reisinger (2012): "Either or Both Compeition: a "two-sided" theory of advertising with overlapping viewerships," Mimeo, Duke University.

Anderson, S. and S. Coate (2005): "Market provision of broadcasting: a Welfare Analysis," *The Review of Economic Studies*, 72, 947-972.

Athey, S.C., E. Calvano and J.S. Gans (2012), "The Impact of the Internet on Advertising Markets for News Media," mimeo, Rotman School of Management, University of Toronto.

Butters, G. (1977): "Equilibrium Distributions of Sales and Advertising Prices," *The Review of Economic Studies*, 44, 465-491.

Appendix

A Proofs

A.1 Proof of Propositions 1 and 2.

We shall split the problem into two subproblems as follows. Problem 1:

$$\max_{n_1,n_2\geq 0} v\left(1-\left(1-\psi\left(n_1\right)\right)\left(1-\psi\left(n_2\right)\right)\right)-p\left(n_1+n_2\right) \quad \text{s.t. } n_1+n_2=n.$$

Let $n_1(n)$ and $n_2(n)=n-n_1(n)$ be the solution of Problem 1. Problem 2:

$$\max_{n\geq 0} v\left(1-\left(1-\psi\left(n_1(n)\right)\right)\left(1-\psi\left(n-n_1(n)\right)\right)\right)-pn.$$

a) Consider the case when $1-\psi$ is log-concave. By the property of log-concave functions for any n_1 and n:

$$\begin{aligned}\left(1-\psi\left(n_1\right)\right)=\left(1-\psi\left(\frac{n_1}{n}n\right)\right)=\left(1-\psi\left(\frac{n_1}{n}n+\left(1-\frac{n_1}{n}\right)0\right)\right)\geq\\ \geq\left(1-\psi\left(n\right)\right)^{\frac{n_1}{n}}\left(1-\psi\left(0\right)\right)^{1-\frac{n_1}{n}}=\left(1-\psi\left(n\right)\right)^{\frac{n_1}{n}}\end{aligned}$$

so:

$$\left(1-\psi\left(n_1\right)\right)\geq\left(1-\psi\left(n\right)\right)^{\frac{n_1}{n}} \tag{16}$$

Similar result holds for n_2:

$$\left(1-\psi\left(n_2\right)\right)\geq\left(1-\psi\left(n\right)\right)^{1-\frac{n_1}{n}} \tag{17}$$

Multiply both parts of (16) by $(1-\psi(n_2))$ (it is strictly positive) and use inequality (17):

$$
\begin{aligned}
&(1-\psi(n_1))(1-\psi(n_2)) \geq \\
&\geq (1-\psi(n))^{\frac{n_1}{n}}(1-\psi(n_2)) \geq (1-\psi(n))^{\frac{n_1}{n}}(1-\psi(n))^{1-\frac{n_1}{n}} = (1-\psi(n_1+n_2))
\end{aligned}
$$

So,

$$(1-\psi(n_1))(1-\psi(n_2)) \geq (1-\psi(n_1+n_2))$$

Multiply by -1 both parts and add 1:

$$1-(1-\psi(n_1))(1-\psi(n_2)) \leq \psi(n_1+n_2)$$

Hence, for any given n full concentration gives higher profit than any level of diversification. So, in the solution either n_1 or n_2 is equal to 0.
Now consider the second problem that in the case of full concentration can be written as:

$$\max_{n\geq 0} v\psi(n) - pn$$

FOC of this problem is:

$$v\psi'(n) = p$$

It has unique solution.
Derive w.r.t. v:

$$\psi'(n) + v\psi''(n)\frac{\partial n}{\partial v} = 0$$

$$\frac{\partial n}{\partial v} = -\frac{\psi'(n)}{v\psi''(n)} > 0$$

Last inequality holds by properties of function ψ.
b) Consider the case of log-linear $1-\psi$. We apply similar argument as in

case a) but for log-linear function. Now:

$$(1-\psi(n_1)) = \left(1-\psi\left(\frac{n_1}{n}n\right)\right) = \left(1-\psi\left(\frac{n_1}{n}n+\left(1-\frac{n_1}{n}\right)0\right)\right) = \\ = (1-\psi(n))^{\frac{n_1}{n}}(1-\psi(0))^{1-\frac{n_1}{n}} = (1-\psi(n))^{\frac{n_1}{n}}$$

So:

$$(1-\psi(n_1)) = (1-\psi(n_1+n_2))^{\frac{n_1}{n}}$$

and

$$(1-\psi(n_2)) = (1-\psi(n_1+n_2))^{1-\frac{n_1}{n}}$$

So, for any n_1 and n_2 we have:

$$1-(1-\psi(n_1))(1-\psi(n_2)) = 1-(1-\psi(n_1+n_2))^{\frac{n_1}{n}}(1-\psi(n_1+n_2))^{1-\frac{n_1}{n}} = \\ = 1-(1-\psi(n_1+n_2)) = \psi(n_1+n_2) \tag{18}$$

Hence, for any fixed n and $n_1 \in [0,n]$ the following equality holds:

$$\psi(n) = 1-(1-\psi(n_1))(1-\psi(n-n_1))$$

The second part of the proof is identical to case a).
c) Consider the case of log-convex $1-\psi$. By the same argument as before:

$$1-(1-\psi(n_1))(1-\psi(n-n_1)) \geq \psi(n)$$

So, for any n and $n_1 \in [0,n]$, any degree of diversification is more optimal than full concentration. Find the optimal degree of diversification for a fixed n:

$$\max_{k\in(0,1)} 1-(1-\psi(kn))(1-\psi((1-k)n))$$

FOC is:

$$(1-\psi((1-k)n))\psi'(kn)=(1-\psi(kn))\psi'((1-k)n)$$

or

$$\frac{\psi'(kn)}{(1-\psi(kn))}=\frac{\psi'((1-k)n)}{(1-\psi((1-k)n))} \tag{19}$$

By properties of log-convex functions, $\frac{\psi'(x)}{1-\psi(x)}$ is monotone increasing in x, so the LHS of (19) is monotone increasing with k when it's RHS is monotone decreasing. Hence it has a unique solution for k, because trivial solution $k=\frac{1}{2}$ always holds, it is unique solution. So maximum diversification is always optimal in this case.

We can rewrite the final problem as follows:

$$\max_{n\geq 0} v\left(1-\left(1-\psi\left(\frac{n}{2}\right)\right)^2\right)-pn$$

FOC is:

$$v\left(1-\psi\left(\frac{n}{2}\right)\right)\psi'\left(\frac{n}{2}\right)=p$$

Derive w.r.t. v:

$$\left(1-\psi\left(\frac{n}{2}\right)\right)\psi'\left(\frac{n}{2}\right)+v\frac{1}{2}\left(\left(1-\psi\left(\frac{n}{2}\right)\right)\psi''\left(\frac{n}{2}\right)-\left(\psi'\left(\frac{n}{2}\right)\right)^2\right)\frac{\partial n}{\partial v}=0$$

$$\frac{\partial n}{\partial v}=-\frac{\left(1-\psi\left(\frac{n}{2}\right)\right)\psi'\left(\frac{n}{2}\right)}{v\frac{1}{2}\left(\left(1-\psi\left(\frac{n}{2}\right)\right)\psi''\left(\frac{n}{2}\right)-\left(\psi'\left(\frac{n}{2}\right)\right)^2\right)}>0$$

by properties of ψ.

Proof of Proposition 4

Start with log-concave case. For any given n in case of full concentration advertiser gets:

$$\pi_C = \left(D^l + D^s\right)\psi\left(n\right)$$

when in case of diversification he gets:

$$\pi_D = 2D^l\psi\left(n\right) + D^s\left(1-\left(1-\psi\left(\frac{n}{2}\right)\right)^2\right)$$

The difference $\pi_C - \pi_D$ is:

$$\begin{aligned}\pi_C - \pi_D &= D^s\left(\psi\left(n\right)-\left(1-\left(1-\psi\left(\frac{n}{2}\right)\right)^2\right)\right) - D^l\psi\left(n\right) = \\ &= D^s\left(\psi\left(n\right)-\left(1-\left(1-\psi\left(\frac{n}{2}\right)\right)^2\right)\right) - \frac{1-D^s}{2}\psi\left(n\right)\end{aligned}$$

where $\psi\left(n\right)-\left(1-\left(1-\psi\left(\frac{n}{2}\right)\right)^2\right) > 0$ for log-concave case, so the difference is linearly increasing with D^s from $-\frac{1}{2}\psi\left(n\right) < 0$ to $\psi\left(n\right)-\left(1-\left(1-\psi\left(\frac{n}{2}\right)\right)^2\right) > 0$. So, for every n there is a point $\tilde{D}^s\left(n\right)$ where he is indifferent between two cases:

$$\tilde{D}^s\left(n\right) = \frac{\frac{1}{2}\psi\left(n\right)}{\frac{3}{2}\psi\left(n\right)-\left(1-\left(1-\psi\left(\frac{n}{2}\right)\right)^2\right)}$$

And for all $D^s < \tilde{D}^s\left(n\right)$ diversification is optimal, when for all $D^s > \tilde{D}\left(n\right)$ concentration is optimal.

In log-linear case:

$$\pi_C - \pi_D = D^s\left(\psi\left(n\right)-\left(1-\left(1-\psi\left(\frac{n}{2}\right)\right)^2\right)\right) - D^l\psi\left(n\right) = -D^l\psi\left(n\right) < 0$$

so, for every n, so diversification always brings higher profit.

Proof of Proposition 5

By Proposition 4, full diversification is always optimal in log-linear case, so advertiser maximizes π_D. But in log-linear case from (18):

$$\pi_D = 2D^l\psi(n) + D^s\left(1 - \left(1 - \psi\left(\frac{n}{2}\right)\right)^2\right) = 2D^l\psi(n) + D^s\psi(n) = \psi(n)$$

that does not depend on D^s.

Proof of Proposition 6

By properties of log-linear function the following equality holds:

$$1 - \left(1 - \psi\left(\frac{\tilde{n}_1}{2}\right)\right)\left(1 - \psi\left(\frac{\tilde{n}_2}{2}\right)\right) = \psi\left(\frac{\tilde{n}_1}{2} + \frac{\tilde{n}_2}{2}\right)$$

so we can rewrite the problem as follows:

$$\max_{\tilde{n}_1,\tilde{n}_2} D_1^l\psi(\tilde{n}_1) + D^s\psi\left(\frac{\tilde{n}_1}{2} + \frac{\tilde{n}_2}{2}\right) - \left(\tilde{n}_1\left(D_1^l + \frac{D^s}{2}\right) + \tilde{n}_2\frac{D^s}{2}\right)\frac{p}{v}$$

FOCs of this problem are:

$$D_1^l\psi'(\tilde{n}_1) + \frac{D^s}{2}\psi'\left(\frac{\tilde{n}_1}{2} + \frac{\tilde{n}_2}{2}\right) = \left(D_1^l + \frac{D^s}{2}\right)\frac{p}{v}$$

$$\frac{D^s}{2}\psi'\left(\frac{\tilde{n}_1}{2} + \frac{\tilde{n}_2}{2}\right) = \frac{D^s}{2}\frac{p}{v}$$

Obviously, $\tilde{n}_1 = \tilde{n}_2$ s.t. $\psi'(\tilde{n}_1) = \frac{p}{v}$ is a solution of this problem. And because ψ is concave, solution is decreasing with $\frac{p}{v}$.

Proof of Proposition 7

In general case, optimal allocation is a solution for:

$$\max_{\tilde{n}_1,\tilde{n}_2} D_1^l\psi\left(\tilde{n}_1\right) + D_2^l\psi\left(\tilde{n}_2\right) +$$
$$+ D^s\left(1-\left(1-\psi\left(\frac{\tilde{n}_1}{2}\right)\right)\left(1-\psi\left(\frac{\tilde{n}_2}{2}\right)\right)\right) - \left(p_1\tilde{n}_1\left(D_1^l+\frac{D^s}{2}\right)+p_2\tilde{n}_2\left(D_2^l+\frac{D^s}{2}\right)\right)\frac{1}{v}$$

where $\tilde{n}_i = \frac{n_i}{D_i^l+\frac{D^s}{2}}$
and k-efficient allocation is a solution for:

$$\min_{\tilde{n}_1,\tilde{n}_2} \tilde{n}_1\left(D_1^l+\frac{D^s}{2}\right)+\tilde{n}_2\left(D_2^l+\frac{D^s}{2}\right)+$$
$$-\left[D_1^l\psi\left(\tilde{n}_1\right) + D_2^l\psi\left(\tilde{n}_2\right) + D^s\left(1-\left(1-\psi\left(\frac{\tilde{n}_1}{2}\right)\right)\left(1-\psi\left(\frac{\tilde{n}_2}{2}\right)\right)\right)\right]$$

$$s.t.\ \ D_1^l\psi\left(\tilde{n}_1\right) + D_2^l\psi\left(\tilde{n}_2\right) + D^s\left(1-\left(1-\psi\left(\frac{\tilde{n}_1}{2}\right)\right)\left(1-\psi\left(\frac{\tilde{n}_2}{2}\right)\right)\right) = k$$

Compare FOCs for optimal allocation:

$$D_1^l\psi'\left(\tilde{n}_1\right)+\frac{D^s}{2}\left(1-\psi\left(\frac{\tilde{n}_2}{2}\right)\right)\psi'\left(\frac{\tilde{n}_1}{2}\right)=\left(D_1^l+\frac{D^s}{2}\right)\frac{p_1}{v}$$

$$D_2^l\psi'\left(\tilde{n}_2\right)+\frac{D^s}{2}\left(1-\psi\left(\frac{\tilde{n}_1}{2}\right)\right)\psi'\left(\frac{\tilde{n}_2}{2}\right)=\left(D_2^l+\frac{D^s}{2}\right)\frac{p_2}{v}$$

and for efficient one:

$$\left(1+\lambda\left(k\right)\right)\left[D_1^l\psi'\left(\tilde{n}_1\right)+\frac{D^s}{2}\left(1-\psi\left(\frac{\tilde{n}_2}{2}\right)\right)\psi'\left(\frac{\tilde{n}_1}{2}\right)\right]=\left(D_1^l+\frac{D^s}{2}\right)$$

$$(1+\lambda(k))\left[D_2^l\psi'(\tilde{n}_2)+\frac{D^s}{2}\left(1-\psi\left(\frac{\tilde{n}_1}{2}\right)\right)\psi'\left(\frac{\tilde{n}_2}{2}\right)\right]=\left(D_2^l+\frac{D^s}{2}\right)$$

For these two systems of equations to coincide we need:

$$\frac{p_1}{v}=\frac{1}{1+\lambda(k)}$$

$$\frac{p_2}{v}=\frac{1}{1+\lambda(k)}$$

or, $p_1 = p_2$. So, for every k there are prices $p_1 = p_2 = \frac{v}{1+\lambda(k)}$, such that solution of the optimal allocation problem coincides with solution of k -efficiency problem.

For every $p_1 = p_2 = p$ we setk^* :

$$k^*=D_1^l\psi(\tilde{n}_1^*)+D_2^l\psi(\tilde{n}_2^*)+D^s\left(1-\left(1-\psi\left(\frac{\tilde{n}_1^*}{2}\right)\right)\left(1-\psi\left(\frac{\tilde{n}_2^*}{2}\right)\right)\right)$$

where$\tilde{n}_1^*$ and $\tilde{n}_2^*$ are solutions of the optimal allocation problem, $\tilde{n}_1^*$ and $\tilde{n}_2^*$ will be also solutions for efficiency problem with k^* . To prove it, suppose per contra that solution of efficiency problem $(\tilde{n}_1^k, \tilde{n}_2^k$) is not equal to $\tilde{n}_1^*, \tilde{n}_2^*$, so we can guarantee k^* with smaller $\tilde{n}_1\left(D_1^l+\frac{D^s}{2}\right)+\tilde{n}_2\left(D_2^l+\frac{D^s}{2}\right)$:

$$\tilde{n}_1^*\left(D_1^l+\frac{D^s}{2}\right)+\tilde{n}_2^*\left(D_2^l+\frac{D^s}{2}\right)>\tilde{n}_1^k\left(D_1^l+\frac{D^s}{2}\right)+\tilde{n}_2^k\left(D_2^l+\frac{D^s}{2}\right)$$

But then, for the value function of optimal allocation π :

$$\pi\left(\tilde{n}_1^*,\tilde{n}_2^*\right)=k^*-\left(\tilde{n}_1^*\left(D_1^l+\frac{D^s}{2}\right)+\tilde{n}_2^*\left(D_2^l+\frac{D^s}{2}\right)\right)\frac{p}{v}<$$
$$<k^*-\left(\tilde{n}_1^k\left(D_1^l+\frac{D^s}{2}\right)+\tilde{n}_2^k\left(D_2^l+\frac{D^s}{2}\right)\right)\frac{p}{v}=\pi\left(\tilde{n}_1^k,\tilde{n}_2^k\right)$$

that contradicts the fact that $\tilde{n}_1^*$ and $\tilde{n}_2^*$ are solutions of this problem. So, $(\tilde{n}_1^k, \tilde{n}_2^k)$ is equal to $(\tilde{n}_1^*, \tilde{n}_2^*)$.

B Instrumental Results (used in proofs)

B.1 Log-linear and log-additive tracking functions

Definition 1. *A positive real valued function $f : X \to \mathbb{R}_+$ is log-concave if $\log f$ is concave, log-convex if $\log f$ is convex and log-linear if $\log f$ is linear.*

Straightforward calculus delivers the following equivalence (proof omitted):

Lemma 2. *Consider a positive, real valued continuous and twice differentiable function $f : X \to \mathbb{R}_+$ The following are equivalent:*

1. *f is log-concave,*
2. *f'/F is monotone decreasing.*

Analogously, log-convexity is equivalent to f'/F increasing and log-linearity to f'/F constant.

An important result which will be key in the first proposition is the following. Suppose $\psi : \mathbb{R}_{++} \to [0,1]$ is continuous, a.e. differentiable, increasing, concave, bounded above by one and satisfies the following boundary conditions: $\psi(0) = 0$, $\lim_{n\to\infty} \psi(n) = 1$.

Lemma 3. *$(1 - \psi(n))$ is log-linear if and only if $(1 - \psi(n_1))(1 - \psi(n_2))$ is constant over the set $\{n_1, n_2 \geq 0 : n_1 + n_2 = n\}$ for all $n \geq 0$.*

Proof. Take any $(\hat{n}_1, \hat{n}_2) \in \{n_1, n_2 > 0 : n_1 + n_2 = n > 0\}$ and let $k > 0$ be the value of $(1 - \psi(n_1))(1 - \psi(n_2))$ computed at that point. Consider the following (non empty) level set $A := \{n_1, n_2 \geq 0 : (1-\psi(n_1)(1-\psi(n_2)) = k\}$. Note that A is symmetric in the sense that if $(n_1 = a, n_2 = b) \in A$ then $(n_1 = b, n_2 = a) \in A$. Strict monotonicity implies that for each n_1 there is at most one n_2 such that $(n_1, n_2) \in A$ which, by continuity and surjectivity of ψ, always exists. Let $n_2(n_1) > 0$ be the implicit function, which is a.e. differentiable with slope equal to

$$-\frac{\psi'(n_1)}{(1 - \psi(n_1))} / \frac{\psi'(n_2(n_1))}{1 - \psi(n_2(n_1))}. \tag{20}$$

By lemma 2, log-linearity of $1 - \psi$ is equivalent to (20) being constant. Symmetry of A and continuity of $n_2(n_1)$ imply that the point $n_1 = n_2$ also belongs to A. (20) computed in $n_1 = n_2$ equals -1, which then implies that $n_2(n_1) = n - n_1$. That is $(\hat{n}_1, \hat{n}_2) \in \{(n_1, n_2) > 0 : n_1 + n_2 = n\}$ implies $(1 - \psi(\hat{n}_1))(1 - \psi(\hat{n}_2)) = k$. $\square$

A second important and, as we shall see, intuitive property of our tracking function is the following:

Definition 2. *A positive real valued function $f : X \to \mathbb{R}_+$ is log-additive if $\log f$ is additive. That is: $\log f(x + y) = \log f(x) + \log f(y)$ for all $x, y \in X$; f is log-superadditive if the former relation holds with $>$ and f is log-subadditive if the former relation holds with $<$.*