Draft model – AC meets GSYY

Justin Katz

July 2022

This document is linked to this Github issue: link. The purpose is mostly on measurement, not analysis of counterfactual equilibria.

1 Perfect targeting within group i

Index advertisers by k so that the value of showing $a_ik \equiv (a_{i1k}, a_{i2k})'$ ads on platforms 1 and 2 respectively is $G_{ik}(a_{ik}, v_k)$ where $v_k \sim F$ and $v_k \equiv (v_{1k}, v_{2k})'$. For posted cost per impression p_{ij} on platform j, advertiser k chooses to buy a_{ij} ads so that:

$$\frac{\partial G_i}{\partial a_{ij}}(a_{ik}) + v_{jk} = p_{ij} \quad \forall j$$

As an aside, it feels like it could be important to also have a shifter v_{ik} since there could be heterogeneity across advertisers in their marginal valuation of a consumer in demographic category i.

Suppose we take a simple parameterization of $\frac{\partial G_i}{\partial a_{ij}}$ by taking the first-order Taylor expansion around zero, or equivalently assume that G_i is quadradic in a_i. Note: we can get similar results if we assume some power function and take logs with multiplicative errors. Then for $j = 1, 2$:

$$p_{ij} = c_{0ij} + g_{i1j} a_{i1k} + g_{i2j} a_{i2k} + v_{kj}$$

Make two assumptions:

- Slutsky symmetry: $g_{i12} = g_{i21}$
- Common intercepts: $c_{0i1} = c_{0i2}$. This might be less reasonable, since the marginal value to the first ad might vary by platform on average. But as we will see this is the "least" of our worries.

Define:

$$g_i \equiv \begin{pmatrix} g_{i11} & g_{i21} \\ g_{i21} & g_{i22} \end{pmatrix}, \quad c_{0i} \equiv \begin{pmatrix} c_{0i1} \\ c_{0i2} \end{pmatrix}, \quad p_i \equiv \begin{pmatrix} p_{i1} \\ p_{i2} \end{pmatrix}$$

Then:

$$p_i = c_{0i} + g_i a_{ik} + v_k$$

If we have data on prices by platform and ad volumes by platform, we could theoretically estimate g_i, c_{0i} from this pricing equation. However, doing so (using e.g. OLS) would require:

$$E\left[a_{ijk} v_{jk}\right] = 0$$

which almost certainly does not hold, since when v_{jk} is high the advertiser will endogenously choose higher a_{ijk}. So we need an instrument for ad demand.

- However, note that once we have (g_i, c_{0i}) we can recover ν_k pointwise.

Can we use equilibrium conditions along with knowledge of time spent on platform to identify g_i, c_{0i}? It turns out we can't, without further assumptions. Let $(\kappa T)_i \equiv (\kappa_{i1}T_{i1}, \kappa_{i2}T_{i2})'$. Then by market clearing:

$$\begin{aligned}(\kappa T)_i &= \int a_{ik} dF(\nu_k) \\ \implies g_i(\kappa T)_i &= \int (p_i - c_{0i} - \nu_k) dF(\nu_k) \\ &= p_i - c_{i0} - \int \nu_k dF(\nu_k)\end{aligned}$$

A reasonable normalization would have $\int \nu_k dF(\nu_k) = 0$, so:

$$p_i = g_i(\kappa T)_i + c_{i0}$$

Intuitively, this just says that the price is the marginal value of the average advertiser, were she to buy all the ad space. This is because we assume the idiosyncratic demand shocks are mean zero and additive.

Suppose we know time spent on platform and can separately estimate κ, and also know prices for showing ads to group i. Then the above is two equations with 6 unknowns (4 parameters in g_i, 2 parameters in c_{i0}). With the Slutsky symmetry and intercept restrictions, it's down to 2 equations with 4 unknowns, which is still not enough. Options:

- Can we use platform profit maximization somehow? Probably not, since we likely need to use that to back out how sensitive consumers are to ads.
- We could make more restrictions – i.e. how much a marginal ad on IG impacts the value of a marginal ad on FB and vice versa.
- We could add another shock so that $p_{it} \neq g_i(\kappa T)_{it} + c_{i0}$, assume that the shock is orthogonal to time spent on platform, and estimate the above using time series regression.

2 Imperfect targeting with overlap

Suppose there is some heterogeneity in platform consumers. Specifically, there are three types of consumers:

- Consumers who only use platform 1 who spend time T^u_{i1} on the platform;
- Consumers who only use platform 2 who spend time T^u_{i2} on the platform;
- Consumers who use both platform 1 and platform 2. They spend time T^b_{i1} on platform 1 and time T^b_{i2} on platform 2.

If there's a unit mass of consumers, and fraction $1-2\mu_i$ are on both platforms, then mechanically a fraction μ_i of consumers are only on either platform 1 or platform 2. Specifically:

$$T_{ij} = \mu_i T^u_{ij} + (1-2\mu_i) T^b_{ij}$$

Advertisers k purchase a mass of ad slots on each platform, $a_{ik} \equiv (a_{i1k}, a_{i2k})$. Purchasing ads is equivalent to purchasing a unit of time on the platform, since supply comes from consumer time use. Advertisers can buy a certain number of impressions on each platform, but not a specific number of impressions per person. This is motivated by institutional relevance and is also convinient:

- First, advertisers buy ads by setting an overall budget for a specific target audience. Advertisers monitor impressions per person, but for default auction buys, (i) Meta only reports averages, not the distribution; and (ii) frequency is someting that advertisers frequently report trying to control by adjusting targeting parameters, budget, creatives, etc.
  - Meta offers reach and frequency ad buying, which allows for buys at a specific frequency, but industry sources say over 95% of ad dollars are spent on auction buys. Reach and frequency buying is typically very expensive and often low quality. Rich Burns mentioned that Facebook buckets users into actions, and so uses people with very predictable click behavior for reach and frequency – these tend to be low quality consumers (though he asked me not to publicize this).
  - Meta also has a “frequency cap” feature. But (i) it only allows caps, not targeting a specific frequency; and (ii) it seems like it is not a very well-known feature – a brand has to call their rep and describe to them how to enable it (see this post: link).
- Second, this assumption makes analysis much more simple, since otherwise we would have to specify a mechanism for rationing. The reason is that if different groups of people on the platform have different “supply” but impressions to both groups cost the same, then it is not possible to clear both markets with one price.

The return that advertiser k gets from showing a minutes of ad time to a given person in demographic group i is $G_{ik}(a)$. Note that relative to the previous section, we assume that ad spend on each platform is perfectly substitutable. We could relax this, but let's roll with it for now.

There are two cases to consider:

- Case 1. Segmented ad markets with no cross-platform tracking. This is probably the more relevant case when platforms are separated. Advertisers place ad buys on each platform and cannot uniformly distribute ad time across platforms.
- Case 2. Single ad market. This might be the more relevant case when platforms are jointly owned (e.g. FB / IG). In this case, the platform pools total ad demand and uniformly distributes across users on both platforms. This is similar to what Meta does when advertisers let the algorithm optimize spend across both platforms.

While the exact empirical implementation of each case is different, they are conceptually pretty similar as I show below.

Case 1. Segmented ad markets without cross-platform tracking For each incremental unit of ad time purchased, the platform randomly selects an incremental unit of ad time produced by consumers in demographic group i. Therefore, the total number of ads served to consumers unique to the platform and who multihome is:

$$E[a^u_{ijk}] = a_{ijk}\mu_i \frac{T^u_{ij}}{T_{ij}}$$

$$E[a^b_{ijk}] = a_{ijk}(1-2\mu_i)\frac{T^b_{ij}}{T_{ij}}$$

This implies the amount of ad time viewed by a member of u or b is $a_{ijk}T^u_{ij}/T_{ij}$ and $a_{ijk}T^b_{ij}/T_{ij}$, respectively. The advertiser problem is therefore:

$$\max_{a_{ik}} \mu_i \sum_{l=1,2} G_{ik}\left(a_{ilk}\frac{T^u_{il}}{T_{il}}\right) + (1-2\mu_i)G_{ik}\left(a_{i1k}\frac{T^b_{i1}}{T_{i1}} + a_{i2k}\frac{T^b_{i2}}{T_{i2}}\right) - (p_{i1}a_{i1k} + p_{i2}a_{i2k})$$

where p_{ij} is the price of ads for demographic group i on platform j. The FOCs help us back out the G_{ik} functions:

$$p_1 = \mu_i \frac{T^u_{i1}}{T_{i1}} G'_{ik}\left(a_{i1k}\frac{T^u_{i1}}{T_{i1}}\right) + (1-2\mu_i)\frac{T^b_{i1}}{T_{i1}} G'_{ik}\left(a_{i1k}\frac{T^b_{i1}}{T_{i1}} + a_{i2k}\frac{T^b_{i2}}{T_{i2}}\right)$$
$$p_2 = \mu_i \frac{T^u_{i2}}{T_{i2}} G'_{ik}\left(a_{i2k}\frac{T^u_{i2}}{T_{i2}}\right) + (1-2\mu_i)\frac{T^b_{i2}}{T_{i2}} G'_{ik}\left(a_{i1k}\frac{T^b_{i1}}{T_{i1}} + a_{i2k}\frac{T^b_{i2}}{T_{i2}}\right)$$

Everything is observable except three objects (with notation $g_{ik} \equiv (g_{ik1}, g_{ik2}, g_{ik3})$:

$$g_{ik1} \equiv G'_{ik}\left(a_{i1k}\frac{T^u_{i1}}{T_{i1}}\right)$$
$$g_{ik2} \equiv G'_{ik}\left(a_{i2k}\frac{T^u_{i2}}{T_{i2}}\right)$$
$$g_{ik3} \equiv G'_{ik}\left(a_{i1k}\frac{T^b_{i1}}{T_{i1}} + a_{i2k}\frac{T^b_{i2}}{T_{i2}}\right)$$

so we have a system with 2 equations and 3 unknowns. This is better than what we had before – with only one restriction we know G_{ik} at three points for each k! The reason is that (i) we assumed that platforms are perfect substitutes; and (ii) we are using a lot of information on audience overlap across platforms. But this gets us a long way:

- *Counterfactual identification*. Although levels aren't identified, changes $g_{ik1} - g_{ik2}$ and $g_{ik2} - g_{ik3}$ are. If G'' is linear, then second derivatives for each advertiser are identified. Counterfactuals mostly depend on diversion ratios, which depend on G'' only (and not G'); we might be able to identify counterfacutal prices even without levels.

- *First-order approximation*. A first-order approximation to G' (which is a second-order approximation to G) relies on two parameters and so would pin down gik. We could test whether this is a good approximation by using the above bullet to see whether $G'' = 0$.

- *Structure on G*. We can get further with more structure on G. For example, suppose we assume that $G_{ik}(\cdot) = G_i(\cdot) + \nu_k$. Parameterize $G_i(\cdot;\theta)$ where θ has N parameters. Suppose there are I demographic groups and K observed advertisers, meaning we have $2KI$ equations and $K+N$ parameters to find. We can parameterize G_i with up to $2K(I-1)$ parameters, which is probably a lot. Different assumptions let us parameterize diminishing returns with varying granularity.

What if platforms are differentiated, i.e. return to a_{ik} is $G_{ik}(a_{i1k}, a_{i2k})$? Then the problem is:

$$\max_{a_{ik}} \mu_i G_{ik}\left(a_{i1k}\frac{T^u_{i1}}{T_{i1}}, 0\right) + G_{ik}\left(0, a_{i2k}\frac{T^u_{i2}}{T_{i2}}\right) + (1-2\mu_i) G_{ik}\left(a_{i1k}\frac{T^b_{i1}}{T_{i1}}, a_{i2k}\frac{T^b_{i2}}{T_{i2}}\right) - (p_{i1}a_{i1k} + p_{i2}a_{i2k})$$

The FOCs are:

$$p_{i1} = \mu_i \frac{T^u_{i1}}{T_{i1}} \frac{\partial G_{ik}}{\partial a_{i1k}}\left[a_{i1k}\frac{T^u_{i1}}{T_{i1}}, 0\right] + (1-2\mu_i)\frac{T^b_{i1}}{T_{i1}} \frac{\partial G_{ik}}{\partial a_{i1k}}\left[a_{i1k}\frac{T^b_{i1}}{T_{i1}}, a_{i2k}\frac{T^b_{i2}}{T_{i2}}\right]$$
$$p_{i2} = \mu_i \frac{T^u_{i2}}{T_{i2}} \frac{\partial G_{ik}}{\partial a_{i2k}}\left[0, a_{i2k}\frac{T^u_{i2}}{T_{i2}}\right] + (1-2\mu_i)\frac{T^b_{i2}}{T_{i2}} \frac{\partial G_{ik}}{\partial a_{i2k}}\left[a_{i1k}\frac{T^b_{i1}}{T_{i1}}, a_{i2k}\frac{T^b_{i2}}{T_{i2}}\right]$$

It's clear there are now four unknowns with two equations and based on how we parameterize G_{ik} we may be able to get a pretty good sense for what diminishing returns look like.

Case 2. Single ad market Now suppose platforms are jointly owned, and the owner chooses to serve ads across platforms as though it is a single market (note: would need to verify that this is optimal). Specifically, for each incremental unit of ad time purchased, the platform selects uniformly a unit of ad time supplied by a consumer in demographic group i from one of the platforms. The combined impression volume purchased by advertisers is pooled across platforms and denoted a^s_{ik} where the s superscript differentiates from the vector from the previous section and stands for single ad market. Let $T_i \equiv T_{i1} + T_{i2}$. For consumers who only use platform j, the expected number of impressions per person is:

$$\frac{T^u_{ij}}{T_i}(a^s_{ik})$$

and for consumers who use both platforms, the epxected number of impressions is:

$$\frac{T^b_{i1} + T^b_{i2}}{T_i}(a^s_{ik})$$

The expected number of ads show on platform j is $T_{ij}/T_i a^s_{ik}$, so the expected cost of bundle a^s_{ik} is:

$$\left(\frac{T_{i1}}{T_i}p_{i1} + \frac{T_{i2}}{T_i}p_{i2}\right)a^s_{ik}$$

so the advertiser problem is:

$$\max_{a^s_{ik}} \mu_i \sum_{l=1,2} G_{ik}\left(\frac{T^u_{il}}{T_i}(a^s_{ik})\right) + (1-2\mu_i)G_{ik}\left(\frac{T^b_{i1} + T^b_{i2}}{T_i}(a^s_{ik})\right) - \left(\frac{T_{i1}}{T_i}p_{i1} + \frac{T_{i2}}{T_i}p_{i2}\right) \cdot a^s_{ik}$$

The FOC has 1 equation and three unknowns (G' evaluated at three points). This requires more restrictions for identification.

Comments:

- We lost information because since we assumed a unified market, the advertisers are now only making one choice, not two.
- The single ad market – where the owner chooses placement uniformly across platforms – is similar to Meta Automatic Ad Placements (which the platform recommends). However, Meta also allows manual placements as well, so this model isn't quite complete. It seems like advertisers can choose to solve the maximization problem in this case or the previous case.
  - If we can observe or infer what brands are doing, then we should incorporate that margin of choice into the model
  - If we observe or infer what brands are doing, then we might be able to use this information to infer substitutiability across platorms, since platforms trade off more efficient targeting in the sense of reducing variation in aggregate G with more efficient targeting in the sense of better targeting impressions x platform.

Next steps

- Market clearing – depends on parameterization of G. This defines a mapping between ad load κ_{ij} and the price vector
- Specify the platform optimization problem – choose κ_{ij} to maximize profits, assuming no consumer response
- Add the consumer response.