The Diffusion of Services

BARAK LIBAI, EITAN MULLER, and RENANA PERES

Web Appendix

Web Appendix A: The solution of Equation 1

The category growth equation (Equation 1) is described by the following:

(A.1) $$\frac{d\,N(t)}{d\,t} = p\big(m - N(t)\big) + \frac{q\,(1-\delta)N(t)}{m}\big(m - N(t)\big) - \delta\,N(t)$$

Denote $q' = q(1-\delta)$. Thus, Equation A.1 is equivalent to:

(A.2) $$\frac{d\,N(t)}{d\,t} = \frac{-q'\,N^2(t)}{m} + (q' - p - \delta)N(t) + p \cdot m$$

The right-hand side of the equation is a quadratic polynomial. Its roots are denoted by r_1 and r_2, and are given by $r_{1,2} = \frac{-\beta \pm \Delta}{\frac{-2q'}{m}} = \frac{\beta \mp \Delta}{\frac{2q'}{m}}$, where $\beta \equiv q' - p - \delta$, and $\Delta \equiv \sqrt{\beta^2 + 4q'p}$.

Recall that for a quadratic equation with roots r_1 and r_2, $Ax^2 + Bx + C = A(x - r_1)(x - r_2)$.
Thus, we can perform a separation of variables and transform the equation to:

(A.3) $$\frac{d\,N}{\left(\frac{q'N}{m} - r_1\frac{q'}{m}\right)(r_2 - N)} = d\,t \quad \Rightarrow \frac{d\,N}{(r_2 - r_1)\left(\frac{q'N}{m} - r_1\frac{q'}{m}\right)} + \frac{d\,N}{\frac{q'}{m}(r_2 - r_1)(r_2 - N)} = dt$$

Integrating Equation A.3 under the initial condition $N(0) = 0$, we get:

(A.4) $$N(t) = \frac{m\dfrac{\Delta+\beta}{2q(1-\delta)}(1-e^{-\Delta\cdot t})}{1+\dfrac{\Delta+\beta}{\Delta-\beta}e^{-\Delta\cdot t}} \qquad \begin{cases}\beta \equiv q(1-\delta)-p-\delta \\ \Delta \equiv \sqrt{\beta^2+4q(1-\delta)p}\end{cases}$$

Defining $\overline{m} = m\dfrac{\Delta+\beta}{2q(1-\delta)}$, $\overline{p} = \dfrac{\Delta-\beta}{2}$, and $\overline{q} = \dfrac{\Delta+\beta}{2}$, and substituting into Equation A.4 yield the solution (i.e., Equation 2).

Web Appendix B: Simulations

To assess the reliability of the category and brand level models, we performed a series of simulations where we generated data with our models, estimated their parameters and observed whether our models return the correct parameters. The parameter estimation was performed on clean data, and on noisy data in several noise levels. Two noise generation mechanisms were used: absolute noise, and relative noise. In absolute noise, each point of data was perturbed with a normally distributed noise term with a fixed standard deviation (normalized as a percentage of market potential). Such a noise enhances the perturbation of the data points with low values, corresponding to early stages in the market growth. In relative noise, we added a noise term to each data point, such that its standard deviation was a certain percentage of the point's original value. This method enhances the noise on the later periods of the market evolution.

Category level model: We generated a full factorial combination of parameter values in typical ranges. We generated data using these parameters and the category level services growth model (Equation 1), and performed nine sets of 64 simulations each– one with clean data, four in various levels of absolute noise (noise levels of 2%, 4%, 6%, and 10%), and four in various levels of relative noise (noise levels of 2%, 4%, 6%, and 10%). Market potential was normalized to the value of 1. The estimations were performed using SAS (proc model, SUR option). To avoid a possible dependency of the estimation on initial parameter values, we used the multiple starting points option, taking a fixed anchoring starting point to be the average values of diffusion parameters of p = 0.03, q = 0.38 (Sultan, Farley and Lehmann 1990), and a typical disadoption value of 0.1. Table B.1 displays the number of estimations which were significantly different from the original parameters at the at the 5% level.

Table B.1: Category level simulations; set size=64 simulations; 192 parameters per set

Noise type	Noise level	No of parameters significantly different from true values	Percentage
no noise	0	0	0
absolute	2%	6	3%
absolute	4%	16	8%
absolute	6%	20	10%

absolute	10%	21	10%
relative	2%	0	0
relative	4%	5	2.6%
relative	6%	14	7%
relative	10%	27	14%

Brand level model: We randomly sampled 300 combinations of a full factorial set of parameter values in typical ranges, generated data with these parameters using the brand level services growth model (Equation 4), and then performed ten sets of 300 simulations each – one with clean data, four with absolute noise (noise levels of 1%, 2%, 4%, and 6%), and five in various levels of relative noise (noise levels of 1%, 2%, 4%, 6%, and 10%). Starting points were chosen as in the category level simulations. The estimations were performed using SAS (proc model, SUR option). Table B. 2 displays the number of estimations which were significantly different from the original parameters at the at the 5% level. The table implies that the number of cases of wrong estimations is small even in the high noise levels. Average adjusted R^2 (averaged over all the simulations) is 78%.

Table B. 2: Brand level simulations; set size=300 simulations; 2400 parameters per set

Noise type	Noise level	No of parameters significantly different from true values	Percentage
no noise	0	0	0
absolute	1%	63	2.6%
absolute	2%	87	3.6%
absolute	4%	102	4.25%
absolute	6%	128	5.3%
relative	1%	9	0.4%
relative	2%	12	0.5%
relative	4%	59	2.4%
relative	6%	95	3.9%
relative	10%	155	6.4%

Web Appendix C: The relationship between the competitive model and the category model

In this Appendix, we investigate the relationship between the competitive model of Equation 4 and the category-level model, described in Equation 1. If one assumes that the attrition levels are the same for all firms (an assumption supported at least by our data), then summing up Equation 4 for all firms, and rearranging terms yield the following equation:

$$\text{(C.1)} \qquad \frac{d\,N(t)}{d\,t} = p(m - N(t)) + \sum_i q_i (1-\delta) N_i(t) \frac{(m - N(t))}{m} - \delta\, N(t)$$

Where $N(t) = \sum_i N_i(t)$, and $p = \sum_i p_i$.

The derivation of (C.1) relies on the fact that since the sum of ε_{ij} for all i should be equal to one, then the following string of equalities holds:

$$c\sum_i \sum_{j \neq i} \varepsilon_{ij} N_j = c\sum_j \sum_{i \neq j} \varepsilon_{ij} N_j = c\sum_j N_j \sum_{i \neq j} \varepsilon_{ij} = c\sum_j N_j = cN$$

Note that we assume that the effective word-of-mouth parameter is reduced by the disadoption rate and not by the overall attrition rate. The reason is that this is the only assumption consistent with the category level, as otherwise the summation of the subscribers of all competing firms does not yield the category level subscribers.

Suppose we were to rewrite Equation 1 but with p and q that are related to the individual firm level by the following: $p = \sum p_i$ and $q = \sum q_i / k$, i.e., the internal parameter q of Equation 1 is the average of the individual firms' internal parameters. One might ask the question about the difference between the two equations (1 and C.1) from a practical point of view. The answer is that the difference is surprisingly small. In comprehensive simulations that we conducted on the numerical solutions of the equations we found that relatively large differences in q_i will lead to small deviations in dN/dt as calculated from Equations 1 and C.1. For example if we take the average values of the parameters in our data set (quarterly data and annual averaged separately) we get that for the annual data, while the ratio of the largest to the smallest q is 2, the percent average deviation of dN/dt as calculated from Equations 1 and C.1 is 18.2%. For the quarterly data the ratio is 2.4 while the percent difference is 13.6%.

Web Appendix D: The solution of Equation 4

What we show in this Appendix is that under some restrictions on the parameters, Equation 4 has a closed form solution. Equation 4 is the following:

(D.1) $$\frac{d\,N_i(t)}{d\,t} = p_i\big(m - N(t)\big) + \frac{q_i\,(1-\delta_i)N_i(t)}{m}\big(m - N(t)\big) - a_i\,N_i(t) + \sum_{j\neq i}\varepsilon_{ij}c_j N_j(t)$$

To solve this equation, what is needed is a term for the total market *N(t)*. The latter is solvable under the following conditions: for all firms *i:*

(D.2) $q_i = q,\ \ a_i = a,\ \ c_i = c,\ \ \varepsilon_i = 1/(k-1)$, where *k* is the number of firms.

One should note that the solution is not symmetric since the coefficients of innovation (p_i) are not the same across the firms. Denote $q' = q(1-\delta)$. Thus, under the above conditions, Equation D.1 is equivalent to:

(D.3) $$\frac{d\,N_i(t)}{d\,t} = p_i\big(m - N(t)\big) + \frac{q'\,N_i(t)}{m}\big(m - N(t)\big) - a\,N_i(t) + \frac{c}{k-1}\sum_{j\neq i}N_j(t)$$

We consider a scenario derived from Equation D.3, where at time $t = 0$, a service firm enters the market. The firm can be either a pioneer or a late entrant, therefore the initial condition for the category is $N(0) = N_0$. The solution for *N(t)* is given by Equation A.4 of Web Appendix A where $a + c = \delta$. Using *N(t)*, Equation D.3 can be integrated. The solution can be presented as follows:

(D.4) $$N_i(t) = \frac{K + Le^{-(\bar{p}+\bar{q})t} - (K+L)e^{-(a+c/(k-1)+\bar{q}-q')t}}{\left(\dfrac{R}{\bar{p}} + \dfrac{\bar{q}}{\bar{m}\,\bar{p}}e^{-(\bar{p}+\bar{q})t}\right)}$$

Where $K = \dfrac{p_i mc + (c/(k-1) - p_i)\bar{m}c}{\bar{p}(a + c/(k-1) + \bar{q} - q')}$, $L = \dfrac{p_i m\bar{q} + (c/(k-1) - p_i)\bar{m}\,\bar{p}}{\bar{m}\,\bar{p}(a + c/(k-1) - q' - \bar{p})}$, $R = \dfrac{N_0\bar{q} + \bar{m}\,\bar{p}}{\bar{m}\,(\bar{m} - N_0)}$,

As in the Bass function, this penetration function is S-shaped, it is zero at time zero, and at infinity, it approaches $K\,\bar{p}/R$.