Critical Mass – a simple operationalization*

Michal Grajek

CIC, Wissenschaftszentrum Berlin and *Humboldt University*

Tobias Kretschmer

ICE, University of Munich and *Centre for Economic Performance, LSE*

Abstract

This paper develops a simple method of identifying the strength of network effects and the degree of compatibility among different versions of a new technology. Further, it provides a rigorous definition of *critical mass*, a concept frequently used in the popular business press. We derive managerial implications of our results and outline the data required to perform accurate predictions based on our model.

Keywords: Critical Mass, Diffusion, Compatibility

JEL Codes: O33, O32

---

* Preliminary version, please do not cite.

1. Introduction

When PayPal went public in 2002, what set it apart from the myriad flops in the dot.com market was that it was considered to have reached *critical mass*.[1] When Joltage attempted to build a WiFi hotspot network, providers of complementary products were hesitant to invest before the technology had reached *critical mass*.[2] The business press carries an abundance of references to the all-important critical mass point at which a technology's success is virtually guaranteed. One of the main difficulties that observers often face however is that critical mass is essentially a retrospective phenomenon – you'll know it when you see it.

Ever since Schelling's (1978, p. 95) assertion that critical mass can be expressed in the form of "critical number, critical density, [and] critical ratio" with the main defining feature being that "some activity is self-sustaining once the measure of that activity passes a certain minimum level," scholars have advanced their own definitions of critical mass points. However, many of these definitions have not been helpful because they were not easily operationalized in an empirical setting. We close this gap by offering a simple, but operationalizable definition of critical mass with the aim of i) supplying practitioners, policymakers and scholars with an econometrically feasible way of identifying and forecasting critical mass, ii) introducing a structural demand model with network effects generating estimates of the strength of network effects and the degree of compatibility and allowing for identification of critical mass, and iii) modelling industries with multiple, incompatible technologies competing for early customers. An added bonus of our paper is that the data requirements are modest – to estimate critical mass, we only require a price index and diffusion figures by supplier.

[1] The Economist, *Party like it's 1999?* 23 February 2002, p. 65.
[2] The Economist, *Making Wi-Fi pay*. 6 June 2002, p. 53.

In the following section, we review the existing literature on critical mass and introduce our structural model of network demand that allows us to recover the strength of network effects and compatibility between technologies as estimated parameters in Section 3. We then illustrate how these parameters can be used to estimate and forecast critical mass and note the managerial and empirical implications of our model.

2. Critical Mass – Previous Literature

The term critical mass has been used in a variety of (both theoretical and empirical) studies. This section reviews the existing literature and the way in which previous articles have used and measured critical mass.

Oren and Smith (1981) and Economides and Himmelberg (1995) develop theoretical models of critical mass, and Economides and Himmelberg (1995) subsequently test it on the US fax machine market. The overall gist of our paper is similar to ours: Assuming a certain structure on demand in a market, they derive results on when critical mass is likely to be reached in such markets. However, Oren and Smith (1981) do not test their model empirically, and Economides and Himmelberg (1995) use their model to estimate the strength of network effects, and their results refer to different market structures and their implications on the equilibrium network size.

Cool et al. (1997) and Loch and Huberman (1999) develop models of the diffusion of technologies within (Cool et al., 1997) and across (Loch and Huberman, 1999) organizations. Their focus, however, is not on estimating a rigorous model of demand displaying critical mass phenomena, but rather on the explanation of a rapid diffusion process, and in particular the period in which diffusion speed is highest, i.e. when a technology seems to "take off". Similarly, Dollinger (1990) develops a model in which collective strategies are explained conceptually, where "collective strategies" are considered to be a process whereby firms in fragmented industries behave alike. The threshold to an emerging collective strategy is effectively the "critical mass" point beyond which a strategy or a pattern of behaviour is adopted by the vast majority of firms in the industry.

Perhaps closest to our work is the theoretical model developed by Cabral (forthcoming), in which Cabral identifies a catastrophe point at which, under certain assumptions on the model parameters, the diffusion process display a discontinuous jump at a certain level of market penetration – the critical mass point.

While there is a general consensus that critical mass clearly reflects a qualitative change in the (aggregate) adoption behaviour of an economic system, identifying it is not easy, especially if done prospectively (rather than retrospectively) and with secondary data (rather than customized data, like company-specific questionnaires). In the model we develop, our aim is to offer a rigorous definition of critical mass in a network market, and to develop a procedure to implement the concept empirically.

3. The Economic Model

3.1. Willingness-to-Pay Function

The demand model we use is a partial equilibrium, discrete choice, dynamic model. The good considered is non-durable, ex ante homogenous[3] and subject to network externalities. We refer to the good supplied by different firms as brands. A consumer's willingness to pay for a given brand is influenced by her type and the network size of that brand. The set of subscribers is referred to as the network. Denote

[3] By *ex ante* homogeneity we mean that different brands of the good are perceived as intrinsically equal. However, the difference in their valuation is possible *ex post* if they have different network sizes.

by $i(i=1,2,...I)$ each brand and assume that there is a measure one of infinitely-lived consumers with unit demand for the good. Consumer *v's* preference for brand i at time t is represented by the willingness-to-pay function $u(v, x_i(t-\delta))$, where v is the individual preference parameter, $x_i(t-\delta)$ is lagged network size of brand i and the perception lag δ a non-negative number.[4] Formally, we assume that the individual preference parameter v is distributed over the interval [0,1] according to the cumulative density function $F(v)$, and that $u(v, x_i(t-\delta))$ is strictly increasing and continuous in v. By construction, the parameter v establishes a rank ordering of consumers in their willingness to pay. We assume that the ranking is invariant with respect to changes in $x_i(t-\delta)$. Without loss of generality we assume also that the higher v, the larger the benefit of using each network. Network effects are captured by the dependence of each consumer's willingness to pay on network size $x_i(t-\delta)$.

The introduction of lagged network size $x_i(t-\delta)$ to the willingness-to-pay function is crucial to our model as an equilibrium selection device to give the unique diffusion path of each network i, although this raises questions about consumers' rationality. Cabral (1990) argues that if the lag δ is infinitely small, consumers are rational because their subscription decisions are identical to the ones done by forward-looking consumers. However, the construction of the willingness-to-pay function does not allow for coordination for a switch from the low-adoption to the high-adoption steady state when both are feasible. Section 3.5 explains these findings in detail.

Empirically, lagged network size in the willingness-to-pay function corresponds to the lagged dependent variable in the estimated equation, making estimation tractable. While it would be preferable to have the actual expectations of consumers, this data is

[4] We expand on the role of δ in Section 3.5.

rarely available. Working with lags represents a sensible alternative, although the approximation of rational consumers' expectations through lags depends on the frequency of our observations.[5]

3.2. Subscription Demand

Each brand i's network consists of a set of subscribers. If brands are incompatible, each makes up its own network so $x_i(t-\delta) = y_i(t-\delta)$, where $y_i(t-\delta)$ stands for cumulative normalized (by overall market size) subscriptions of brand i. However, if brands are perfectly compatible, the network is common, which is given by total sales of all brands $x_i(t-\delta) = \sum_{j=1}^{I} y_j(t-\delta)$. Brands with identical network size are perceived as perfect substitutes by consumers.

In a more general setting with partial compatibility, brand network size is the weighted sum of its own and all other subscribers. Assuming a symmetric degree of compatibility, we can write this as

$$(1) \qquad x_i(t-\delta) = y_i(t-\delta) + w\sum_{j\neq i} y_j(t-\delta),$$

where $w \in [0,1]$ measures the degree of compatibility, and $w=1$ and $w=0$ correspond to perfect compatibility and perfect incompatibility, respectively. At each time t, consumer v decides to buy one of the brands or to stay out of the market to maximize her net utility

$$(2) \qquad u(v, x_i(t-\delta)) - p_i(t)$$

If (2) is negative for all brands, she will not join any network.[6]

[5] Clearly, the less frequent the observations, the worse the approximation.
[6] This "static" decision rule in our dynamic model is appropriate for non-durable goods.

Denote by $v_{i,t}^{*} = v^{*}(x_i(t-\delta), p_i(t))$ the type of the consumer indifferent between joining network i or not with respect to brand i at time t. $v_{i,t}^{*}$. Rewriting, we get

(3) $$u(v_{i,t}^{*}, x_i(t-\delta)) = p_i(t).$$

The brand i for which $v_{i.t}^{*}$ is the lowest is the most attractive brand for all subscribers at time t. Define

(4) $$v_{L,t}^{*} = \min_i \{v_{1,t}^{*}, v_{2,t}^{*}, \ldots, v_{I,t}^{*}\}.$$

All consumers with a higher preference than $v_{L,t}^{*}$ buy the good. If $v_{i,t}^{*}$ is the same for multiple brands, the subscribers choose among them with equal probability. Define

(5) $$H_i(v_t^{*}) \equiv \begin{cases} \dfrac{1-F(v_{L,t}^{*})}{I_{L,t}} & if \quad v_{i,t}^{*} = v_{L,t}^{*} \\ 0 & otherwise \end{cases},$$

where $v_t^{*} = (v_{1,t}^{*}, v_{2,t}^{*}, \ldots, v_{I,t}^{*})$ is a vector of the indifferent types with respect to brand i at time t, $I_{L,t}$ is the number of brands for which $v_{i,t}^{*} = v_{L,t}^{*}$ and F is the distribution function of v. H_i equals the number of the consumers willing to buy brand i in time t. The state equations describing the evolution of each brand's sales over time are then given by

(6) $$y_i(t) = H_i(v_t^{*}).$$

In the steady state (assuming constant prices) we expect that no consumer can increase her utility by switching brands, so each brand's sales stay constant over time:

(7) $$y_i(t) = y_i(t-\delta).$$

It is worth noting that the steady-state equilibrium demand of the above model coincides with the standard static model equilibrium with fulfilled consumers' expectations (see Rohlfs, 1974, and Economides and Himmelberg, 1995). Moreover, the process of achieving equilibrium we have described formally corresponds to the logic presented by Rohlfs (1974).

3.3. Switching Costs

The above model of demand with network externalities is an extension of Cabral's (1990) single brand model. It possesses however some specific features. One of them is persistent firm symmetry. In each instance of time, every active firm has an equal number of subscribers $y_i(t)$, which stays in contrast to the observation that real firms' market shares exhibit persistent differences. The other feature resembles the Bertrand paradox. If one firm undercuts the others minutely it takes the entire market, which results in fierce price competition. Moreover, with incompatibility it is extremely difficult to recoup market shares once a firm has lost ground, because it needs to offer far more attractive price then its rival with an installed base advantage. Switching costs mediate this tendency of network markets somewhat as market share is more "sticky" if existing consumers find it costly to switch suppliers (Farrell and Klemperer, 2001).

Suppose switching costs are high enough so that, having chosen one particular brand, a consumer would not switch later on, for example because of penalties for cancelling their subscription prematurely, as observed in telecommunications markets, for example. For tractability of our model, however, we assume that switching is possible without costs if prices increase. Such switching costs change the demand described in (5) in that only unattached consumers can join a network. We can rewrite (5) as

$$(8) \qquad H_i^{'}\left(v_t^*, v_{t-\delta}^*\right) \equiv \begin{cases} \dfrac{F(v_{L,t-\delta}^*) - F(v_{L,t}^*)}{I_{L,t}} & if \qquad v_{i,t}^* = v_{L,t}^* \\ 0 & otherwise \end{cases}.$$

Now, H_i' equals the number of the *new* consumers willing to buy the brand i at time t. Consequently, the state equations are given by

$$(9) \qquad y_i(t) = H_i^{'}\left(v_t^*, v_{t-\delta}^*\right) + y_i(t-\delta).$$

In contrast to equation (6), equation (9) exhibits substantial inertia. Any asymmetry inherited from previous periods will not disappear. In particular, an incumbent's installed base will constitute a persistent competitive advantage over an entrant. Substituting definition (6) into (9) and rearranging the terms yields[7]

(10) $y_{i,t} = H_i(v_t*) + l_iE_t,$

where the l_i's are operator-specific constants and E_t is an entry indicator function that is zero pre-entry and one post-entry. One appealing feature of (10) is that it nests two regimes: with and without switching costs. The formal derivation of (10)—shown in the appendix—requires, however, additional constraints on the operators' pricing behaviour. Intuitively, if entry is to be the only event that breaks symmetry, as indicated in (10), we need to assume that each operator strategically adjusts its price in response to changes in own and competitors' installed bases in order to win new subscribers in each period. We cover pricing issues in the next subsection.

3.4. Supply of the Network Good

To complete the model of the market we would need to model how the prices are determined. Since the paper focuses on the demand side of the market and, in particular, on identification of the network effects, we do not introduce a structure for

[7] See appendix for details. For simplicity, we consider single entry only. It is straightforward to extend the model to account for multiple entries.

the supply side. Instead, we discuss some possibilities of extending the economic model to contain the explicit pricing relation as well.

In the simplest case without switching costs, we could plausibly assume that fierce price competition between symmetric firms drives prices down to (common) marginal costs. Price changes over time will then reflect technological progress and/or economies of scale and the market outcome will be completely symmetric. In other words, expression (2) will be equal for all *i*. No firm will drop out of the market $\left(I_{L,t} = I\right)$ and all networks will grow (or decline) equally fast.

With switching costs, a space for strategic pricing emerges. Indeed, switching costs in our set up tend to reduce competition and give firms the opportunity to earn profits. In that case, firms face a trade-off. On the one hand, they want to keep prices high in order to exploit the installed base of consumers. On the other hand, they want to lower prices to attract new subscribers, i.e. to enhance the installed base in the future. Static, marginal-costs pricing is no longer appropriate for modeling this kind of pricing behavior, as it does not take account of this trade-off. Instead, state-space games, in which actions taken in one period shift payoffs in subsequent periods, could be utilized with installed bases of firms as natural state variables (see Basar and Olsder, 1999). Examples of empirical dynamic pricing models within this framework have been developed in the learning-by-doing literature (e.g. Jarmin, 1994). In these models, cumulative past sales benefit firms in that they lower production costs, which gives rise to a similar trade-off as with network externalities and switching costs.

From an econometric perspective, we do not need any structure for the supply relation to be able to correctly estimate the network effect parameter, and endogeneity issues regarding the price variable can be resolved by instrumental variable techniques.

3.5. Dynamics of the Network Good Adoption

In this section, we abstract from switching costs and assume the price to be equal across brands and constant over time, and networks to be compatible for ease of exposition.[8, 9] In other words, there is one network and one competitive price. The evolution of the installed bases of particular brands is therefore proportional to common network evolution. We also assume that the cumulative density function $F(v)$ and the willingness-to-pay function $u(v, x(t-\delta))$ are continuously differentiable in all arguments. A detailed mathematical treatment of the equilibrium network size path in such model can be found in Cabral (1990), who shows that if network effects are sufficiently strong and the lag length δ tends to zero the equilibrium adoption path is unique and discontinuous. The equilibrium adoption path is described by equation (6). Given the assumptions made above, (6) simplifies to

$$(11) \qquad x(t) = H(v_t^*) \equiv 1 - F(v_t^*)$$

To see how network externalities and price affect diffusion we calculate the derivatives of H with respect to the lagged network size $x(t-\delta)$ and price p in the Appendix. Since H maps the network size from time $t-\delta$ to t, it is convenient to think of it as of a function of the lagged network size $x(t-\delta)$. An examination of (A.4) in the appendix leads to the following lemma:

Lemma 1*: Whenever a solution to equation (3) exists so that v_t^* is defined, and the density $f(v_t^*)$ is strictly positive, the extent of network externalities measured by $\frac{\partial u(v_t^*, x(t-\delta))}{\partial x(t-\delta)}$ determine the slope of the function H in the $x(t-\delta)$ domain, such that (i) H is non-decreasing if and only if network externalities are non-negative, (ii)*

[8] A more general graphical analysis is available from the authors on request.
[9] The subscripts *i* are omitted throughout this section.

the slope of H equals zero if there are no network externalities, and (iii) the slope of H is larger if network externalities are stronger, other things being equal.

Figure 1 illustrates the dynamics of the network good adoption. In the top panel we draw H as a function of the lagged network size $x(t-\delta)$. Lemma 1 formalizes the link between the extent of network externalities and the slope of H, which in turn determines diffusion dynamics. The lower panel shows the steady-state equilibria of network size for each price p, denoted $D(p)$. As mentioned above, this coincides with the analogous static model equilibria with fulfilled consumer expectations. Our dynamic model allows, however, for discrimination among multiple steady-state equilibria. Suppose that market price is p^* in Figure 1. According to (9) then, network size will evolve as indicated in the top panel. If it starts at some size smaller than x_1 it will eventually reach x_0, if the starting network size is bigger than x_1 it will end up in x_2. If the network size equals exactly x_1, it will stay there, but any arbitrarily small shock will lead to an equilibrium at x_0 or x_2. Therefore, x_0 and x_2 are stable steady states, while x_1 is unstable. The following Lemma applies this logic to all prices p:

Lemma 2: *Whenever a solution to equation (3) exists so that v_t^* is defined, and the density $f\left(v_t^*\right)$ is strictly positive, changes in price p determine the shifts of the function H in the $x(t-\delta)$ domain, such that $H\left(v^*(x(t-\delta), p_1)\right) > H\left(v^*(x(t-\delta), p_2)\right)$ for every $x(t-\delta)$ if $p_1 < p_2$.*

Lemma 2 follows directly from examination of (A.7) in the Appendix. It says that lowering the price shifts the function H upwards, although not necessarily in a parallel shift. Drawing steady states for each price yields the steady-state demand $D_{(p)}$. We can conclude that:

Theorem 1: *Downward-sloping parts of the steady-state demand* $D_{(p)}$ *consist of stable equilibria, while the upward-sloping parts are unstable, i.e. consist of critical-mass points.*

Now, consider a case when price changes over time. To see how the common network evolves let the price $p(t)$ be a continuous and decreasing function of time and let $p(0) > p_h$ (as in Figure 1) and $x(p(0))$ be the unique steady-state network size given $p(0)$. As price falls, network size follows the lower steady-state size. Eventually the price reaches p_l and just after that the network size jumps discontinuously to the higher steady-state and grows further on along it. Formally, this diffusion pattern has been shown to be correct for infinitely small δ in Cabral (1990). If the perception lag is strictly positive, consumers are myopic with respect to network size; they do not recognize that the network is going to grow in the current period. As a consequence, the equilibrium network size does not follow exactly but rather tends to the steady-state size. There is no discontinuous jump in the network diffusion either, and diffusion follows an S-shape.

This dynamic perspective helps understand the equilibrium selection rule assumed implicitly in our model by the lag structure, since it does not allow for coordination among consumers. Note that it would be Pareto-optimal to jump to the larger steady-state network size before price falls below p_l. However, this would require the coordination of consumers' subscription decisions to reach critical mass. Another insight from the analysis is the substitutability between the effects of the strength of network effects and lag length on diffusion, which is given in the following Theorem:

Theorem 2: *Network externalities measured by* $\frac{\partial u(v_t^*, x(t-\delta))}{\partial x(t-\delta)}$ *and the perception lag* δ *are substitutes in that both strengthening of network externalities and shortening of the lag length speed up adoption of the network good.*

The arrows in top panel of Figure 1 indicate the change in network size from $t-\delta$ to time t. The length of these arrows reflects the speed of network size growth (or decline). Now, strengthening the network effects, which implies a larger slope of H (Lemma 1), and lowering lag length (say to $\delta/2$, so between time $t-\delta$ and t there are two "updates" of network size) one can achieve the same network size growth. In other words, a strong network externalities combined with a large perception lag may result in the same diffusion speed as weak network externalities and a small lag. One should keep in mind this substitutability when interpreting empirical results. Conversely, manipulating the lag does not influence the steady-state equilibria (the fixed points of the function H), while strengthening of network externalities does. This will be important for the empirical identification of the network effects' strength.

4. The Stochastic Model

4.1. An Example of Functional Specification

The next step towards the structural econometric model is to specify functional forms for the underlying economic model. The specification in this section has been chosen for two reasons. First, the specification yields the demand relation as a simple linear equation (in parameters), which is convenient to work with empirically. Second, the demand relation nests the well-established Bass (1969) diffusion model. Assume consumers' willingness-to-pay function to be

(12) $u(v,x_{i,t-1}) = v + cx_{i,t-1} + dx^2_{i,t-1}$,

where c and d are parameters determining the extent of network effects,[10] and the squared term capturing possible nonlinearities, e.g. diminishing marginal network effects usually assumed in the theoretical literature (Swann, 2002). Consumer types v are uniformly distributed over $(-\infty, a]$ with density $b > 0$. For convenience, population size is not normalized to one as in the theoretical model. In fact, given the distribution of types specified above, the population is infinite to avoid corner solutions.[11]

Given these functional forms, the diffusion equation (10) becomes

$$(13) \qquad y_{i,t} = l_i E_t + \frac{1}{I_t}\left(ab - bp_{i,t} + bcx_{i,t-1} + bdx_{i,t-1}^2\right).$$

Finally, to obtain the estimation equation, we substitute (1) into (13)

$$(14) \quad y_{i,t} = \lambda_i E_t + \frac{1}{I_t}\left(\alpha + \beta p_{i,t} + \gamma_1 y_{i,t-1} + \gamma_2 y_{i,t-1}^2 + \gamma_{11} y_{-i,t-1} + \gamma_{21} y_{i,t-1} y_{-i,t-1} + \gamma_{22} y_{-i,t-1}^2\right),$$

where $y_{-i,t-1}$ denotes the sum of subscribers to all operators other than i in period t-1.

4.2. Identification

Our structure corresponds to the information diffusion models widely studied in marketing. In particular, equation (14) nests the original diffusion equation proposed by Bass (1969) for the single product case. For single brand diffusion, (14) simplifies to the original Bass model if $\beta = 0$ (i.e. price does not matter for network diffusion).

The structural parameters of the model can be identified from the coefficients in (14). Simple algebra yields the highest consumer type for the population $a = -\alpha/\beta$ and density of the distribution of types $b = -\beta$. The parameter α can be interpreted as the

[10] Note that in the empirical model δ will be determined by the data frequency; consequently, we replace δ with 1 meaning "one period" from now on.
[11] Alternatively, the distribution support could be bounded from below to limit the population of consumers and the bound assumed to be low enough in order to avoid the necessity of considering corner solutions, when all consumers subscribe.

number of consumers with positive valuation of the mobile telephone service given zero network size. The network effects parameters c and d are - γ_1/β and - γ_2/β, respectively. The compatibility parameter w is overidentified, since it is identified by γ_{11}/γ_1, $\gamma_{21}/2\gamma_2$, and $\pm\sqrt{\gamma_{22}/\gamma_2}$. When network effects are only present at the operator level (i.e. completely incompatible networks, $w = 0$), the coefficients γ_{11}, γ_{21} and γ_{22} equal zero. Conversely, when network effects operate at the industry level (i.e. fully compatible networks, $w = 1$), the following equalities hold: $\gamma_{11} = \gamma_1$, $\gamma_{22} = \gamma_2$ and $\gamma_{21} = 2\gamma_2$. All intermediate cases with partial compatibility can be easily obtained from the three identifying equalities as well.

The interpretation of the identified structural network externalities parameters c and d directly is difficult for two reasons. First, as already indicated in section 4.1, the empirical identification of the network effects relies heavily on functional forms. In particular, the distribution of types plays a key role. Another assumption that affects the estimates of network effects is the consumer perception lag δ imposed by the data frequency. Further, even if we have statistically significant estimates of c and d we still do not know what constitutes economically significant network effects.

Going back to the first problem, our empirical estimates of the network effect can be biased because of the functional assumptions. In particular, the uniform distribution of types is likely to bias the network estimates upward, i.e. to attach a significant proportion of the S-shape to the network effect arbitrarily. The natural assumption is that the distribution of types mimics the distribution of consumer income, which is usually log-normal. Section 3.5 on the dynamics of network growth helps understand how any bell shaped distribution of types generates an s-shaped diffusion curve.[12]

[12] This is the rank effect outlined in Stoneman (2002).

In the case of the perception lag the direction of the bias is less clear. In general, we would expect that imposing larger (smaller) lag than the actual one creates an upward (downward) bias. But the question of the actual lag size remains open. In section 3.5 we formalized the intuitive relationship between the lag length and the strength of network effects. We noted also that the steady-state equilibria are not affected by different lag lengths.

Aware of the possible bias in our estimates, how can we infer the economic significance of the identified network effects parameters? We propose calculating the steady-state inverse demand functions from (14), replacing all the parameters with their empirical estimates and imposing the steady-state conditions (7). All the important economic phenomena driven by network effects, like multiple equilibria and critical mass of adopters, apply to the case with upward sloping demand. Therefore, the existence of an upward sloping part in the empirical steady-state demand function indicates strong network externalities and possible critical mass.

The empirical steady-state demand function also seems more robust to the incorrect functional assumptions than the identified structural parameters themselves. First, the steady-state equilibria are not affected by different lag lengths. Second, intuition suggests that attaching some distribution-of-types effects to network effects should not change the shape of the steady-state demand function dramatically. Therefore, it seems unlikely that we obtain an upward-sloping part of the demand function in the estimation if the true market demand function is downward-sloping all along.

Some additional information about the source of network effects in the market under consideration can be obtained from the estimate of the compatibility parameter w. For example, the ability to satisfy more communication needs with a larger installed base may give rise to direct network externalities which operate at the industry level (for compatible networks). Conversely, endogenous externalities (Blonski, 2002) created by firms charging an access fee for calls from the outside into their network operate at the firm level (implying incompatible brands).

4.4. Simulations

This section provides some numerical examples of the steady state in our theoretical model by taking different values of the structural parameters. For expositional ease, we assume that providers are symmetric and prices across providers will be equal and can be represented by a common market price p. In Figure 2, we draw different steady-state (inverse) demand functions by varying the network effects parameters c and d. The other parameters of the model are set as follows: a = 100 (maximum willingness to pay for zero network size), b = 0.02 (price sensitivity of demand), w = 0.5 (compatibility), and I = 2 (number of firms). With network effects becoming stronger (i.e. increasing c) the demand function becomes more concave and eventually features an upward-sloping part (critical mass).

Figure 3 illustrates how the demand for a network product is sensitive to the degree of compatibility between different brands of the good. The remaining parameters are set as follows: a = 100, b = 0.02, c = 60, and d = -0.5. We see that increasing the compatibility parameter has a similar effect to strengthening the network effect. In particular, it can lead to critical mass phenomena. Increasing the number of (symmetric) firms in turn has a similar effect to decreasing compatibility.

5. Discussion: Adoption Dynamics of Network Goods

In summary, we find the results that i) diffusion is increasing in network effects and decreasing in price, ii) network effects and perception lags are substitutes in terms of diffusion speed, iii) an upward-sloping part of the demand curve defines critical mass points, and iv) the marginal effect of price changes is different pre- and post critical mass. This carries a number of implications for firms operating in network markets. First, large perception lags may slow down the diffusion of a technology. If firms can reduce this lag, for example by publishing past sales figures, this can be a strategic variable that can accelerate market success without having to lower prices radically. This may be particularly profitable for products with strong network effect which might otherwise reach critical mass rapidly, but are slowed down by significant perception lags. Second, the effect of price changes in network markets can differ dramatically depending on whether critical mass has been reached or not. Specifically, lowering prices early on may result in an immediate switch from the low- to the high-adoption equilibrium (a catastrophe point in Cabral's (1990) terminology), while at later stages of diffusion, price changes will result in a more conventional shift of the demand curve and consequently a faster convergence to the steady (high-adoption) state. Finally, our relatively simple model allows for an identification of critical mass using a small number of parameters. Given current prices, a firm can therefore predict how close they are to reaching critical mass and analyze the effect of a change in their strategic variables, for example price and perception lag, as illustrated above. Parameters can either originate from previous demand estimation in a different geographical, but similar economic market,[13] prior market or technical knowledge about the product in question, or from estimation results at an early stage of diffusion. Although the estimated parameters may be somewhat imprecise, a sensitivity analysis may still generate useful results about the product's position on the diffusion curve and the likely effect of strategic changes.

[13] See, e.g. Grajek and Kretschmer (2006) for a multi-country study of the cellular telephony market in which some countries have introduced

6. Conclusion

The simple model we propose captures a number of relevant features of dynamic markets with network effects. The feature frequently missing in existing models was an econometrically feasible, practically relevant technique of estimating critical mass points and their potential consequences. Our model hopes to fill this gap in the literature and provide researchers and practitioners alike with a starting point for further work on industries featuring critical mass.

Appendix

A.1. Derivatives of the function H with respect to x(t-δ) and p

First, note that v_t* is an implicit function of *x(t-δ)* and *p*, what under simplifying assumptions in the section 2.5 is described by

(A.1) $u(v_t*, x(t-\delta)) = p$.

To calculate the derivative of *H* with respect to the lagged network size *x(t-δ)* we first apply the chain rule to the definition of *H* given in (9). We obtain

(A.2) $$\frac{\partial}{\partial x(t-\delta)} H(v_t*) = -\frac{\partial F(v_t*)}{\partial v_t*} \cdot \frac{\partial v_t*}{\partial x(t-\delta)}$$

The first term on the RHS of (A.2) is just the density of *v* at v_t*. To calculate the second term note that the total derivative of $u(v_t*, x(t-\delta))$ with respect to *x(t- δ)* must stay constant in order to satisfy equation (A.1). This holds for

(A.3) $$\frac{\partial u(v_t*, x(t-\delta))}{\partial x(t-\delta)} = -\frac{\partial u(v_t*, x(t-\delta)}{\partial v_t*} \cdot \frac{\partial v_t*}{\partial x(t-\delta)}$$

Solving (A.3) for $\frac{\partial v_t*}{\partial x(t-\delta)}$ and substituting that into (A.2) yields the desired result

(A.4) $$\frac{\partial}{\partial x(t-\delta)} H(v_t*) = f(v_t*) \cdot \left(\frac{\partial u(v_t*, x(t-\delta)}{\partial v_t*}\right)^{-1} \cdot \frac{\partial u(v_t*, x(t-\delta))}{\partial x(t-\delta)},$$

where *f* is a density function of *v*.

Analogously, to calculate the derivative of *H* with respect to the price *p* we first apply the chain rule to obtain

(A.5) $$\frac{\partial}{\partial p} H(v_t*) = -\frac{\partial F(v_t*)}{\partial v_t*} \cdot \frac{\partial v_t*}{\partial p}.$$

Then we note that

(A.6) $$\frac{\partial u(v_t*, x(t-\delta)}{\partial v_t*} \cdot \frac{\partial v_t*}{\partial p} = 1,$$

and substitute to get

(A.7) $$\frac{\partial}{\partial p} H(v_t*) = -f(v_t*) \cdot \left(\frac{\partial u(v_t*, x(t-\delta))}{\partial v_t*}\right)^{-1}.$$

A.2. State equations with Firm-Specific Constants

To see how we can nest the two regimes (with and without switching costs) in a single set of the state equations rewrite (8) using the definitions (4) and (7) to

$$\text{(A.8)} \quad y_i(t) = H_i(v_t^*) + y_i(t\text{-}\delta) - H_i(v_{t\text{-}\delta}^*)\frac{I_{L,t-\delta}}{I_{L,t}}.$$

Remember that we assume equal hedonic prices among firms. You can think of (A.8) as of a decomposition of the total sales of bran i in time t under switching costs. The first term on the RHS of (A.8) gives the total sales of brand i (number of subscribers) if there were no switching costs. The second and the third term adds and subtracts the installed base of brand i respectively in a way that is sensitive to the number of active firms in the market. To see how this can lead to persistent asymmetries among firms expand the recursive equation (A.8) to

$$\text{(A.9)} \quad y_i(T) = H_i(v_T^*) + H_i(v_{T\text{-}\delta}^*) - H_i(v_{T\text{-}\delta}^*)\frac{I_{L,T-\delta}}{I_{L,T}} + H_i(v_{T\text{-}2\delta}^*) - H_i(v_{T\text{-}2\delta}^*)\frac{I_{L,T-2\delta}}{I_{L,T-\delta}} +$$

$$\ldots + y_i(0) - H_i(v_0^*)\frac{I_{L,0}}{I_{L,\delta}},$$

where $t = 0$ indicates the time when the market starts up so there are no sales at that time and $T > 0$.

Suppose, there is constant number of firms active in the market such that $I_{L,t} = I_L$ for $t \in (0,T)$. Then the last two terms on the RHS of (A.9) equal zero, because every firm is active from the very beginning of the market, and all the middle terms cancel out. In this case (A.9) simplifies to (5), i.e. the state equations with and without switching costs are the same.

Now suppose, there was an entry into the market in $t = E$, and $0 < E < T$. This means that $I_{L,t}$ rises discontinuously in $t = E$ and stays at the higher level afterwards. The sales equations of the incumbents do not simplify to (5) any longer. They become instead

$$\text{(A.10)} \quad y_i^{inc}(T) = H_i(v_T^*) + \int_E^{E+\delta} [H_i(v_{t\text{-}\delta}^*) - H_i(v_{t\text{-}\delta}^*)\frac{I_{L,t-\delta}}{I_{L,t}}]dt,$$

for $T \geq E+\delta$. The integral in (A.10) is positive. It is also invariant with respect to any events in $T > E+\delta$ and can be trated therefor as a firm-specific constant in the post-entry period.

In contrast, the expansion of the recursive equation (A.8) does not go back to $t = 0$ for the entrants. Their history starts at $t = E$ and the sales can be described by

$$\text{(A.11)} \quad y_i^{ent}(T) = H_i(v_T^*) + \int_E^{E+\delta} [-H_i(v_{t\text{-}\delta}^*)\frac{I_{L,t-\delta}}{I_{L,t}}]dt,$$

for $T \geq E+\delta$. To see this result, refer to (A.8) and note that $y_i^{ent}(t\text{-}\delta) = 0$ for $t \in (E, E+\delta)$. The integral in (A.11) plays analogous role for the entrants as the integral in (A.10) for incumbents, but it is negative. Therefore, we can conclude that the incumbents have a constant (in terms of the difference in the total sales) competitive advantage over the entrants.

Moreover one can show that the the fixed effects caused by entry sum up to zero. To see that denote the number of incumbents as *A* and the number of entrants as *B*. The sum of the effects is then

$$A \int_E^{E+\delta} [H_i(v_{t-\delta}{}^*) - H_i(v_{t-\delta}{}^*)\frac{I_{L,t-\delta}}{I_{L,t}}]dt + B \int_E^{E+\delta} [-H_i(v_{t-\delta}{}^*)\frac{I_{L,t-\delta}}{I_{L,t}}]dt =$$

$$\text{(A.12)} \quad = A \int_E^{E+\delta} [H_i(v_{t-\delta}{}^*) - H_i(v_{t-\delta}{}^*)\frac{A}{A+B}]dt + B \int_E^{E+\delta} [-H_i(v_{t-\delta}{}^*)\frac{A}{A+B}]dt =$$

$$= \left(A - \frac{A^2}{A+B} - \frac{AB}{A+B}\right) \int_E^{E+\delta} H_i(v_{t-\delta}{}^*)\, dt = 0.$$

One could also investigate the effects of exit in the analogous manner. Since in our economic structure there is no reason for a firm to leave the market we skip this discussion.

Figure 1. Stable vs. unstable equilibria

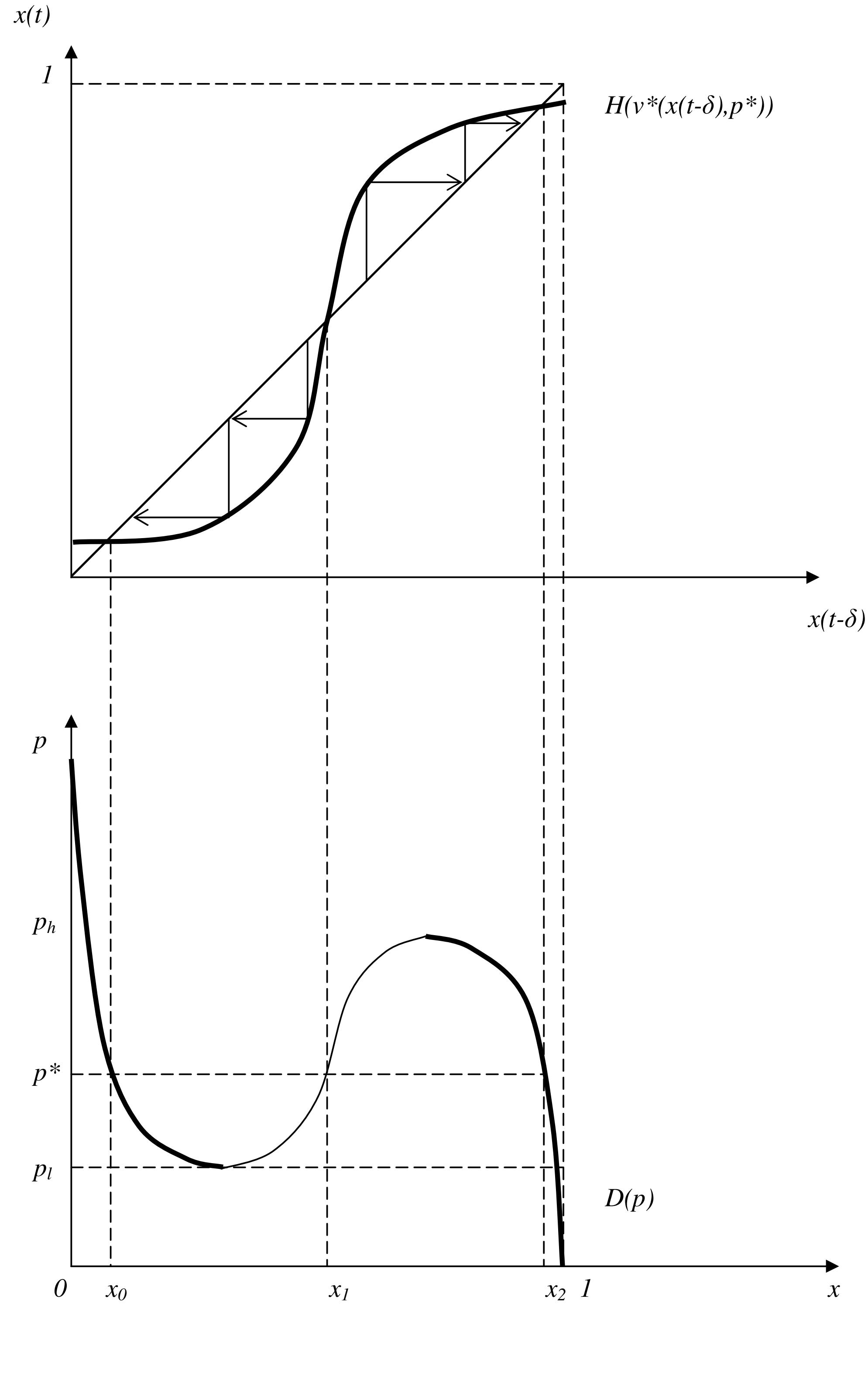


Figure 2. Simulations of the steady-state demand using different values for the network effects parameters.

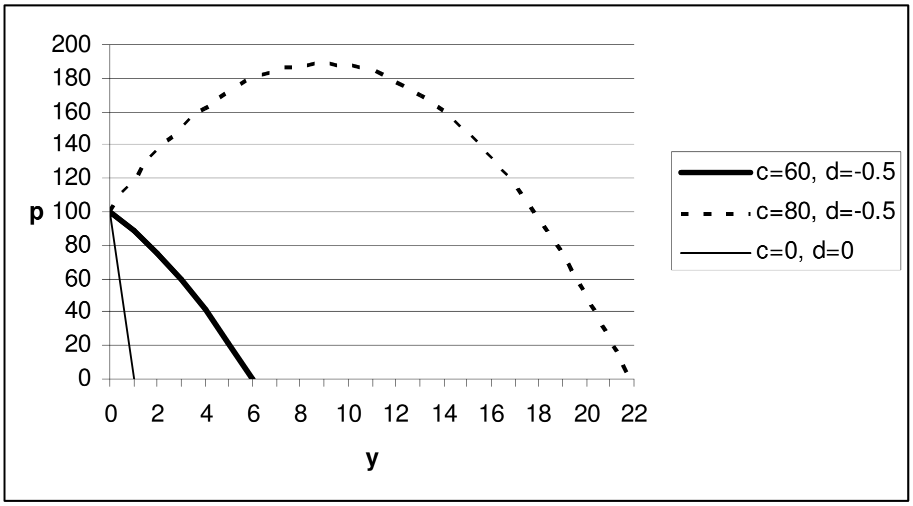
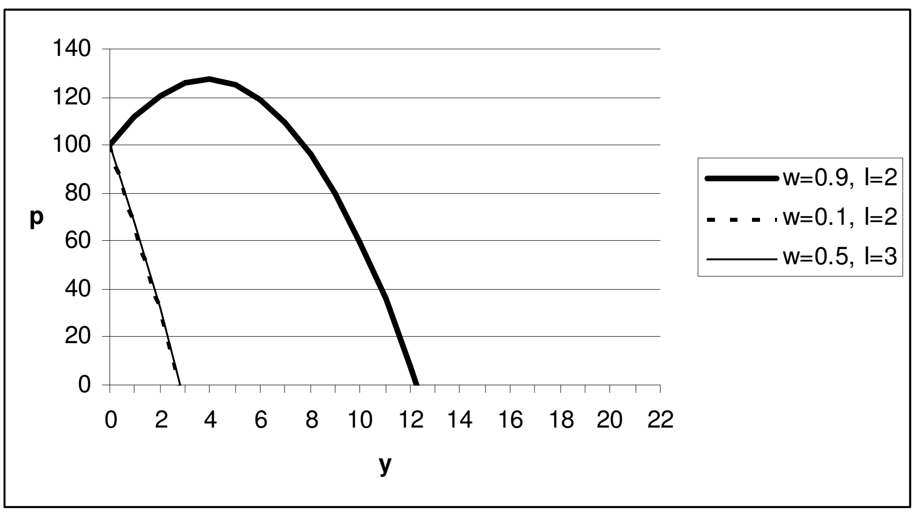


Figure 3. Simulations of the steady-state demand using different values for the number-of-firms and compatibility parameters.



References [PRELIMINARY]

Basar T., G.J. Olsder (1999), “Dynamic Noncooperative Game Theory”, *SIAM*, Philadelphia.

Bass F.M. (1969), “A new product growth model for consumer durables”, *Management Science* 15(5), 215-227.

Blonski M. (2002), “Network externalities and two-part tariffs in telecommunication markets”, *Information Economics and Policy* 14, 95-109.

Cabral L.M.B. (1990), “On the adoption of innovations with ‘network’ externalities”, *Mathematical Social Sciences* 19, 299-308.

Economides N., Ch. Himmelberg (1995), “Critical Mass and Network Size with Application to the US Fax market”, Stern School of Business, New York University *Discussion Paper* EC-95-11.

Farrell J., P. Klemperer (2001), “Coordination and Lock-In: Competition with Switching Costs and Network Effects”, mimeo.

Granger C.W.J., T. Teräsvirta (1993), “Modelling Nonlinear Economic Relationships”, *Oxford University Press*, New York.

Horsky D. (1990), “A Diffusion Model Incorporating Product Benefits, Price, Income and Information”, *Marketing Science* 9(4), 342-365.

Jain D.C., R.C. Rao (1990), “Effect of Price on the Demand for Durables: Modeling, Estimation, and Findings”, *Journal of Business & Economic Statistics* 8(2), 163-170.

Jarmin R.S. (1994), “Learning by Doing and Competition in the Early Rayon Industry”, *Rand Journal of Economics*, 25, 441-454.

Leibenstein H. (1950), “Bandwagon, snob, and Veblen Effects in the theory of consumers’ demand”, *Quarterly Journal of Economics* 64, 183-207.

Rohlfs J. (1974), “A theory of interdependent demand for a communications service”, *Bell Journal of Economics* 5(1), 16-37.