The Evolution of Two-Sided Markets: A Dynamic Model

Feng Zhu
and
Michael Mitzenmacher

TR-02-10

The Evolution of Two-Sided Markets: A Dynamic Model

Feng Zhu and Michael Mitzenmacher[*]

February 7, 2008

Abstract

Static models of two-sided markets often lead to multiple equilibria as a result of increasing return to demand. Typically the market can be monopolistic or oligopolistic in equilibrium. We provide a simple dynamic model to address the equilibrium selection problem. Our results indicate that the final market structure depends critically on the overall strength of the indirect network effects of the two sides. Interestingly, we find under weak network effects the market share advantage of the leader diminishes over time. Simulating our model under the monopoly conditions also suggests that time until dominance follows a power law distribution, and as a result the probability that monopoly does not occur for an extended period is not trivial. We then discuss the implications of our results for intermediaries.

1 Introduction

Many markets in the modern economy involve two groups of participants who need to interact via intermediaries, where an intermediary can be for example a technological standard or a platform. Such markets are known as two-sided markets (e.g. Evans 2003). An example of a two-sided market is video game consoles. The participants are the game developers and the game buyers; the intermediary firms include Sony, Nintendo, and Microsoft, each of which produces (incompatible) game consoles, with each console having its own associated developer and user communities. Other examples include payment cards (card holders and merchants), online auction houses (sellers and buyers), media players (content providers and users), dating clubs (men and women) and operating systems (application developers and users). We henceforth refer to the two sides of the market as buyers and sellers, and the combination of an intermediary with its buyers and sellers as a network.

In these two-sided markets, one side of a network becomes more attractive as the number on the other side increases, a property often referred to as indirect network effects (e.g. Katz and Shapiro 1985). That is, the more buyers, the more attractive the intermediary is for new sellers; similarly, the more sellers, the more attractive the intermediary is for new buyers. Indirect network effects often cause a market to exhibit a rich-gets-richer phenomenon — when a network obtains non-negligible advantages in the market share on both sides, it rapidly grows to a monopoly or near-monopoly. As a result of this process, a particular technological standard or platform provider achieves dominance. For example, the market for VCRs tipped in favor of the VHS standard, and the market for online auction tipped in favor of eBay. However, such two-sided markets do not always tip. Multiple platforms and standards can co-exist, each with a significant market share. This is the case for video game systems as well as payment cards.

Our work is motivated by these observations. We are interested in circumstances under which a two-sided market will ultimately settle on a dominant technological standard or platform, and the amount of time it takes for a dominator to emerge. We are also interested in the likelihood that a follower can catch up and overtake the leader in these markets. In addition, indirect network effects give rise to a "chicken and egg" problem (see, for example, Caillaud and Jullien 2003): to attract more buyers (respectively sellers), an intermediary needs to have a large base of sellers (respectively buyers), but building such a base requires a large based of buyers (respectively sellers) already! Therefore, in-

termediaries need to decide how to allocate their resources to achieve the optimal growth rates of their networks: should they spend most of their resources attracting buyers or attracting sellers? We would like to understand the optimal strategy for a new intermediary in a competitive environment to scale up its network. These questions are critical for the survival of intermediaries. If it appears that a market will tip and the time-to-dominance is short, an intermediary needs to gain advantages in market share in the very beginning in order to succeed.

We develop a probabilistic model based on indirect network effects. We consider a market in which there are two competing networks. In our model, new buyers and sellers arrive asynchronously and they choose a particular network with a probability that depends on the distribution of the existing agents on the other side of the market. We measure the strength of indirect network effects by *market share elasticity of demand* with *market share* defined by the distribution of agents on the other side of the market. We also allow the strength of indirect network effects to differ on both sides of the market.

In this paper, we focus on using our model to derive results on the final state of the market. Our results indicate that the ultimate market structure depends on the product of the market share elasticities of the two sides. When the product is greater than one, one network will eventually become a monopoly, in the sense that the market share of one network will (with high probability) keep increasing and approaching 1. When the product is 1, each network converges to some constant fraction of the market share, depending on the initial state. When the product is less than one the market tends to a state where sellers and buyers are equal for both networks. We also demonstrate heuristics under our model for determining the time until one network dominates the market and for simulating behaviors from various initial conditions.

2   Related Literature

Static game-theoretical models on network effects and two-sided markets (e.g. Armstrong forthcoming, Caillaud and Jullien 2003, Chakravorti and Roson 2004, Chou and Shy 1993, Church and Gandal 1992, Evans 2003, Evans et al. 2004, Farrell and Saloner 1985, 1986, Gabszewicz and Wauthy 2004, Guthrie and Wright 2003, Katz and Shapiro 1985, Rochet and Tirole 2003) often lead to multiple equilibria as a result of the increasing return to demand in these markets. In most cases, in equilibrium the market can be monopolistic or oligopolistic. These models do not address the equilibrium selection problem. In ad-

dition, as most models assume that agents join the market simultaneously, they do not consider the amount of time it takes to reach an equilibrium. Our work is related to literature on evolutionary game theory (e.g. Auriol and Benaïm 2000, Fudenberg and Levine 1998, Kandori et al. 1993, Weibull 1995), which aims at studying equilibrium selection in games that have multiple equilibria. In our probabilistic model we allow agents to join the market asynchronously. Following the treatment in Auriol and Benaïm (2000) and Kandori et al. (1993), we also assume the agents are myopic. That is, agents select networks solely based on current state variables and are not forward-looking. We then illustrate how successive adoption choices made by these myopic agents on both sides eventually aggregate into a collective choice.

Our work is also closely related to probabilistic models of positive feedback (e.g. Arthur 1989, 1994, Arthur et al. 1983). In these models, competing standards and platforms are represented by bins and individual choices of buyers correspond to balls being put into the bins.

Drinea et al. (2002) and Khanin and Khanin (2001) study an specific example of the balls-and-bins process in which a ball joins bin i with probability $\frac{x_i^p}{\sum_{j=1}^n x_j^p}$, where x_i is the number of existing balls in bin i and p is a small positive number. Mitzenmacher et al. (2004) and Oliveira (2004), building on an "exponential embedding" technique first suggested by Rubin (Davis 1990), did an extensive study of this example and also extend some of the results to a more general scenario in which the probability is specified by $\frac{f(x_i)}{\sum_{j=1}^n f(x_j)}$.

One of the main results in this setting shows that the above family of processes undergoes a phase transition at $p = 1$. For $p > 1$ and arbitrary initial conditions, there is one bin that receives all but a finite number of balls in a large-time limit. When $p < 1$ the ratio between the numbers of balls in any pair of bins converges to 1.

In a similar vein, Keilbach and Posch (1998) models the dynamic of technology adoption by $D_k(s_k^t) = (s_k^t)^\lambda p_k^\alpha$, where k refers a particular technology, D_k is the demand of technology k, s_k^t is the market share of technology k at time t and p_k is the price of the technology k. It is easy to see that $\alpha = \frac{\Delta D_k/D_k}{\Delta p_k/p_k}$ and $\lambda = \frac{\Delta D_k/D_k}{\Delta s_k/s_k}$. Therefore, α here is the price elasticity of demand and similarly λ can be called the *market share* elasticity of demand. Although their model also incorporates pricing strategies, their model is equivalent to the ones proposed by the above authors in expectation. If we consider technologies as bins and assume technology i has x_i number of balls, ignoring pricing strategies, Keibach and Posch's model gives $D_i = (\frac{x_i}{\sum_{j=1}^n x_j})^\lambda$. Hence the percentage of new balls that join bin i,

$\frac{D_i}{\sum_{j=1}^n D_j}$, is $\frac{x_i^\lambda}{\sum_{j=1}^n x_j^\lambda}$.

These probabilistic models are suitable for markets with direct network effects, for example, in the case of instant messengers. Many important industries in the modern economy, however, are two-sided and are characterized by indirect network effects. We extend these probabilistic models to study the dynamics of two-sided markets.

3 Model

We consider two two-sided networks, A and B, each consisting of an intermediary, along with groups of buyers and sellers. We assume the two intermediary technologies are identical but incompatible with each other. Each seller sells one type of product. Therefore the utility of an agent, either a buyer or seller, depends only on the relative number of agents on the other side of the same network.

At each stage, a group of N agents join the market. The probability that the group of agents are buyers is α and sellers is $1 - \alpha$. We use α to control for relative arrival rates of the two sides. Each agent joins only one network. We denote the state of network A at time t by the ordered pair $(b_A(t), s_A(t))$ where $b_A(t)$ is the number of buyers in network A and $s_A(t)$ is the number of sellers in network A. We similarly use $(b_B(t), s_B(t))$ for the state of network B. We drop the explicit dependence on t and use b_A, b_B, s_A and s_B when the meaning is clear.

Following Auriol and Benaïm (2000), Nair et al. (2004) and Kandori et al. (1993), we assume that agents are myopic. Their perception of network benefits is based solely on the current state variables and do not consider the impact of their adoption choices on future adopters. At each stage, a new agent selects a network to maximize his short term benefit. This assumption is a good approximation for peer-to-peer markets such as online auction houses and flea markets. It is also applicable to markets in which future benefit is relatively small. For instance, in the case of video game industry the popularity of game titles decreases rapidly after their releases. In addition, new-generation game consoles are introduced every few years. As a result, game developers heavily discount their revenues from future game players. The assumption of myopic agents allows us to consider each period independent of future periods.

We model the competition between two networks using a linear location model. The linear location model has been frequently used to model competition in studies on network effects and two-sided markets (e.g. Armstrong forthcoming, Auriol and Benaïm

2000, Church and Gandal 1992, Clements and Ohashi 2005).

Consider the moment at which a buyer is deciding between two networks. Assume the preference of a buyer for the two networks is characterized by his type δ, which is uniformly distributed in $[0, 1]$. Also assume that network A and network B are located at 0 and 1 respectively. The δ (respectively $1 - \delta$) term can then be interpreted as the "distance" between network A (respectively network B) and the buyer's ideal network, which would have been located at δ. We represent the preferences of a buyer of type δ over network A using the utility function $(1 - \delta)u_A$ and network B using the utility function δu_B. These utility functions capture the idea that the further a network is from the buyer, the less preferred the network. As we are modeling the situation where a buyer's preference for network A or B depends on the number of sellers in the same network, we let $u_A = c_s s_A^p$ and $u_B = c_s s_B^p$. Here p is a small positive number and c_s is a positive constant. Previous studies (e.g. Armstrong forthcoming, Rochet and Tirole 2003) often assume utility functions to be linear in the number of agents for analytical convenience. We use a more flexible functional form to capture linear, convex and concave cases. The buyer prefers the network that provides the higher utility. Letting the location of the indifferent buyer, for whom the utility of network A and B are the same, be δ^*, we have $(1 - \delta^*)u_A = \delta^* u_B$. Solving the equation, we find $\delta^* = \frac{s_A^p}{s_A^p + s_B^p}$. The probability that an incoming buyer joins network A is therefore the probability that $\delta < \delta^*$. As we have assumed δ is uniform on $[0, 1]$, this probability is $\frac{s_A^p}{s_A^p + s_B^p}$. A new buyer joins network B with probability $s_B^p/(s_A^p + s_B^p)$. Similarly, a new seller joins network A with probability $b_A^q/(b_A^q + b_B^q)$ and network B with probability $b_B^q/(b_A^q + b_B^q)$. q is also a small number.

The model indicates that for both kinds of two-sided markets the probability for a randomly selected agent to join a particular network is determined by the distribution of agents on the other side at each stage. We choose to use probabilistic representations as they eliminate potential unstable evolutionary equilibria. The functional forms are similar to those used in Drinea et al. (2002) and Khanin and Khanin (2001). As in Oliveira (2004) and other works, more general functional forms can also be considered. For simplicity, in our analysis we use $N = 1$ per time step and generally set the time parameter t so that it is equal to the total number of buyers and sellers in the system. Figure 1 illustrates our model.

Similar to Keilbach and Posch (1998), p and q can be interpreted *market share elasticities of demand*, except that *market share* here is defined by the number of agents on the other side of the network. It is easy to see that the greater the value of p or q, the more likely the

network with a larger market share on the other side will be selected. Therefore, p and q here indicate the strength of indirect network effects. In addition, as $p, q > 0$, an agent is always more likely to choose the network with a larger market share on the other side. That is, the indirect network effects in our model are always positive.

We ignore strategies by the intermediaries in our model. As a result, the indirect network effects provide the sole motivation for an agent to join a particular network. Our purpose here is to illustrate the evolution path of two-sided markets under indirect network effects. Intermediaries can then adopt various strategies such as advertising and product differentiation to alter the evolution path. In Section 6, we briefly discuss the implication of our results for intermediaries.
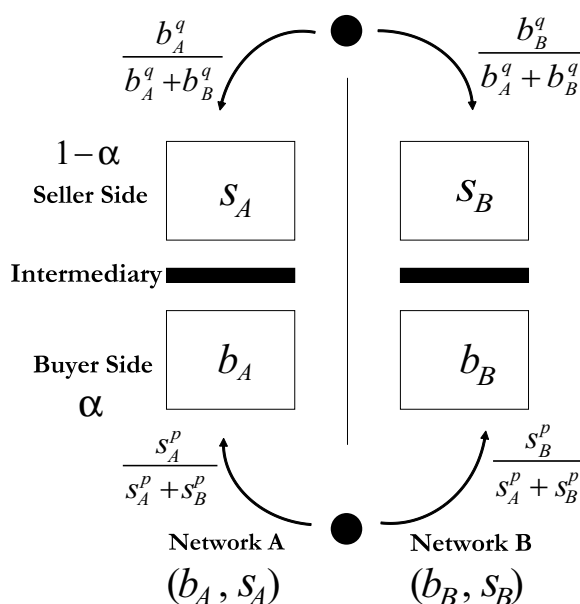

Figure 1: Our probabilistic model. At each step, one agent joins the market. The agent is a buyer with probability α and is a seller with probability $1 - \alpha$. The new agent selects a particular network with probability that depends on the distribution of the existing agents on the other side of the market.

4   A Heuristic Analysis

4.1   Differential Equation Approximation

We start by considering a heuristic approach, following Auriol and Benaïm (2000), Benaïm and Weibull (2003) and Drinea et al. (2002). Then the expected change in $b_A(t)$ over one time step, which we denote by $\Delta b_A(t)$, satisfies

$$\Delta b_A(t) = E[b_A(t+1) - b_A(t)] = \frac{\alpha s_A(t)^p}{s_A(t)^p + s_B(t)^p}. \tag{4.1}$$

Using the heuristic approximation $\Delta b_A(t) = \frac{db_A}{dt}$ and dropping the t from the notation where there is no confusion, we have

$$\frac{db_A}{dt} = \frac{\alpha s_A^p}{s_A^p + s_B^p} \tag{4.2}$$

We can derive similar expressions for $b_B(t)$, $s_A(t)$ and $s_B(t)$, yielding a system of four differential equations, that we refer to henceforth as System A:

$$\frac{db_A}{dt} = \frac{\alpha s_A^p}{s_A^p + s_B^p}, \qquad\qquad \frac{db_B}{dt} = \frac{\alpha s_B^p}{s_A^p + s_B^p},$$

$$\frac{ds_A}{dt} = \frac{(1-\alpha)b_A^q}{b_A^q + b_B^q}, \qquad\qquad \frac{ds_B}{dt} = \frac{(1-\alpha)b_B^q}{b_A^q + b_B^q}.$$

The heuristic yields good approximations of the behaviors of the underlying probabilistic system. This can be formalized for finite time intervals using the theory of martingales; see Kurtz (1981), Mitzenmacher (1996), and Wormald (1995) for related work.

4.2   Alternative representations

The equations in System A dependent on seven parameters: α, p, q, and the initial values of $b_A, b_B, s_A,$ and s_B. We can simplify the setup in various ways that prove useful for reasoning about the system.

For example, we may remove the dependence on t and consider the derived equations:

$$\frac{db_A}{db_B} = \left(\frac{s_A}{s_B}\right)^p, \qquad\qquad \frac{ds_A}{ds_B} = \left(\frac{b_A}{b_B}\right)^q.$$

This representation is useful for considering the fixed trajectories of System A, which correspond to ratios of b_A/b_B and s_A/s_B that remain fixed over time. For example, for any values of p and q the initial conditions where $s_A = s_B$ and $b_A = b_B$ give a fixed trajectory where $s_A = s_B$ and $b_A = b_B$ for all time. As we shall see, this trajectory is not necessarily stable, in the sense that the probabilistic system is likely to deviate from this trajectory rapidly for certain values of p and q.

There may be other fixed trajectories when $pq = 1$. That is, we seek constants β and γ with $\beta = s_A/s_B$ and $\gamma = b_A/b_B$ satisfying

$$\frac{db_A}{db_B} = \left(\frac{s_A}{s_B}\right)^p = \beta^p = \gamma$$

and

$$\frac{ds_A}{ds_B} = \left(\frac{b_A}{b_B}\right)^q = \gamma^q = \beta.$$

This required that $\beta^{pq} = \beta$, or that $pq = 1$ (when $\beta \neq 0, 1$). Any values β and γ satisfying the above equations give a fixed trajectory.

Similar insight can be obtained by letting $r_s = s_A/s_B$ and $r_b = b_A/b_B$. Then

$$\frac{dr_s}{dt} = \frac{(s_B)(ds_A/dt) - (s_A)(ds_B/dt)}{s_B^2} = \frac{(1-\alpha)(r_b^q - r_s)}{s_B(1+r_b^q)},$$

and similarly

$$\frac{dr_b}{dt} = \frac{\alpha(r_s^p - r_b)}{b_B(1+r_s^p)}.$$

While these equations look somewhat less pleasant, they are perhaps more informative. Again, we see that when $pq = 1$ there is now a fixed point for these equations when $r_b^q = r_s$. Moreover, by considering the signs of $\frac{dr_s}{dt}$ and $\frac{dr_b}{dt}$, we can see that there are essentially three different cases, corresponding to the state diagrams given in Figure 2:

Case 1: $pq > 1$. In this case, $(r_s, r_b) = (1, 1)$ is an unstable fixed point; trajectories naturally converge to (∞, ∞) or $(0, 0)$. That is, a monopoly occurs.

Case 2: $pq = 1$. In this case, there is a family of fixed points, and the deterministic system eventually hits one.

Case 3: $pq < 1$. In this case, $(r_s, r_b) = (1, 1)$ is a stable fixed point; trajectories naturally converge so that the system eventually reaches a state where sellers and buyers are

equal for both networks.

The system undergoes a phase transition at $pq = 1$. Note that in all three cases, we are considering positive network effects in the sense that the network with a large market share on the other side is preferred. When $pq > 1$, we would observe the rich-gets-richer phenomena as a result of strong network effects. When $pq \leq 1$, although positive network effects still present, they are not strong enough to drive monopolies to emerge.

Finally, additional insight can be gained by supposing that we start initially with $b_A + b_B = \alpha t$ and $s_A + s_B = (1 - \alpha)t$. Little is lost by this assumption, since from any starting state we will quickly converge to a state where $b_A + b_B \approx \alpha t$ and $s_A + s_B \approx (1 - \alpha)t$. Now suppose that we let

$$b_A(t) = \frac{(1 + \epsilon_b(t))\alpha t}{2}, \qquad\qquad b_B(t) = \frac{(1 - \epsilon_b(t))\alpha t}{2},$$

$$s_A(t) = \frac{(1 + \epsilon_s(t))(1 - \alpha)t}{2}, \qquad\qquad s_B(t) = \frac{(1 - \epsilon_s(t))(1 - \alpha)t}{2}.$$

Now

$$\epsilon_b(t) = \frac{b_A - b_B}{\alpha t}, \qquad\qquad \epsilon_s(t) = \frac{s_A - s_B}{(1 - \alpha)t}.$$

We can therefore examine the "gaps" $\epsilon_b(t)$ and $\epsilon_s(t)$ to gain insight into the rate of convergence.

4.3 Uses of differential equations

The differential equations framework can be used to give insight into the time until one network obtains dominance or the effect of changing the initial conditions. One can simply evaluate the process given by the differential equations numerically. Using the differential equations is generally much faster and simpler than simulating the model directly, especially as one may want to simulate multiple times to account for variations in the random experiment.

For example, starting from the state $(1200, 800)$ for network A and $(1000, 1000)$ for network B, with parameters $p = q = 2.0$ and $\alpha = 2/3$, our simulation of the differential equations indicates that it takes $32, 149$ steps (or equivalently, 32,149 entrants join the

$$r_s^p = r_b \qquad r_b^q = r_s \qquad r_b^q = r_s \quad (\text{or } r_s^p = r_b) \qquad r_b^q = r_s \qquad r_s^p = r_b$$
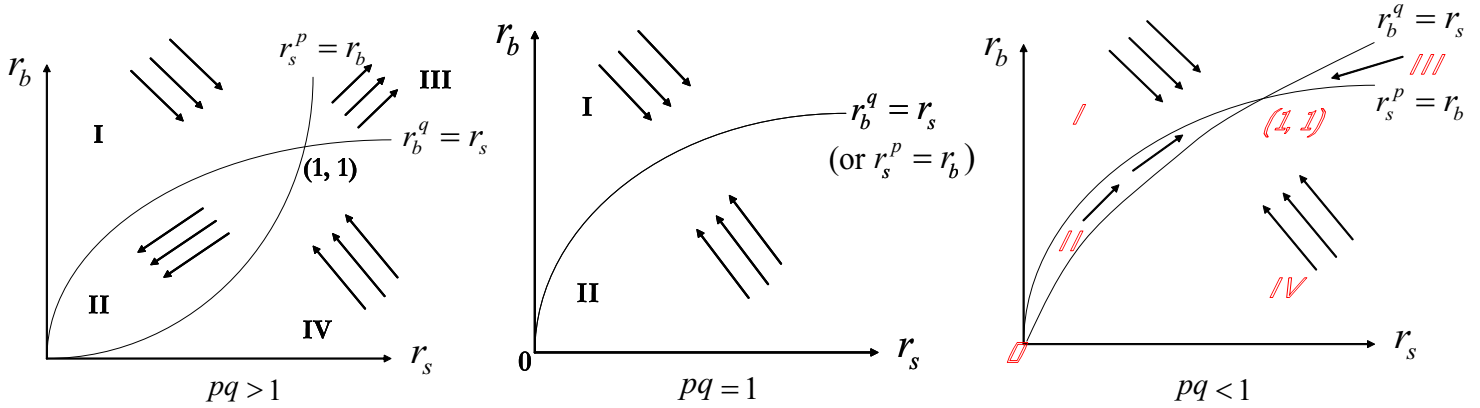
Figure 2: State diagrams for $pq > 1$, $pq = 1$ and $pq < 1$. Arrows are used to indicate the directions of trajectories for states in each region, as determined by signs of $\frac{dr_s}{dt}$ and $\frac{dr_b}{dt}$. Note that the concavity or convexity of the curves can be different from the ones drawn above. For example, in the case where $pq > 1$, both curves can be concave.

system) before network B achieves 60% of the buyers and sellers. We also simulate the stochastic process 20,000 times. The number of steps until network B achieves 60% of the buyers and sellers varies from 10,247 to 263,636,885 with a median value of 32,279 (we excluded the 69 runs in which network A eventually dominates). Note that the median is close to the value predicted by the differential equations, but the mean is significantly larger (just over 65,000 in this case). We discuss this further momentarily.

To see the effects of changing the initial state, we could then examine how the results would change if network B instead started at $(900, 1100)$. The differential equations indicate that it takes 22,906 steps before network B achieves 60% of the buyers and sellers. (While it may not appear intuitive that switching the initial state of network B in this fashion would lessen the time to dominance, recall that buyers are arriving twice as fast as sellers; hence having more sellers initially enhances the advantage for network B.) In this case, network A has never achieved dominance among the 20,000 simulations. The number of steps until dominance ranged from 8,830 to 3,236,117, with a median of 23,054.

We reiterate that the differential equations are only accurate approximations for sufficiently large systems. Thus, while they can give some very useful insight into the behavior of the system the results should be considered carefully in practice. Specifically, in this instance the system is initially sufficiently small that there is some probability that network A emerges dominant, a fact not predictable from the differential equations. Further, another experimental finding not given by the differential equations is that the time

11

until dominance is reached has very large variance. This is not surprising, given the work by Oliveira (2004) in the balls-and-bins setting showing that the time until dominance is achieved essentially follows a power law distribution. A similar effect appears here. We illustrate it with results from the above simulations. We rank the number of steps until dominance and plot the log of the number of steps against the log of the ranks in Figure 3. The relationships are nearly linear on the log-log plots. The tail ends of the distributions are skewed as they represent rare events and we only run a finite number of simulations. The R-squared values from linear regressions are 0.94 and 0.96 respectively. The linear relationships in the log-log plots give evidence that the time until dominance follows a power law distribution.

Oliveira's proof relies largely on the exponential embedding technique, which we have not been able to generalize to this setting. The economic implications of this finding, however, are important: in this setting, while the differential equations accurately predict the *median* time until dominance occurs, the *mean* time is significantly larger, and the probability that dominance does not occur for an extended period is non-trivial.

Although the differential equation approach is heuristic, provable statements based on the discrete probabilistic model can be made, as we demonstrate in the appendix for the case of $p > 1$ and $q > 1$.
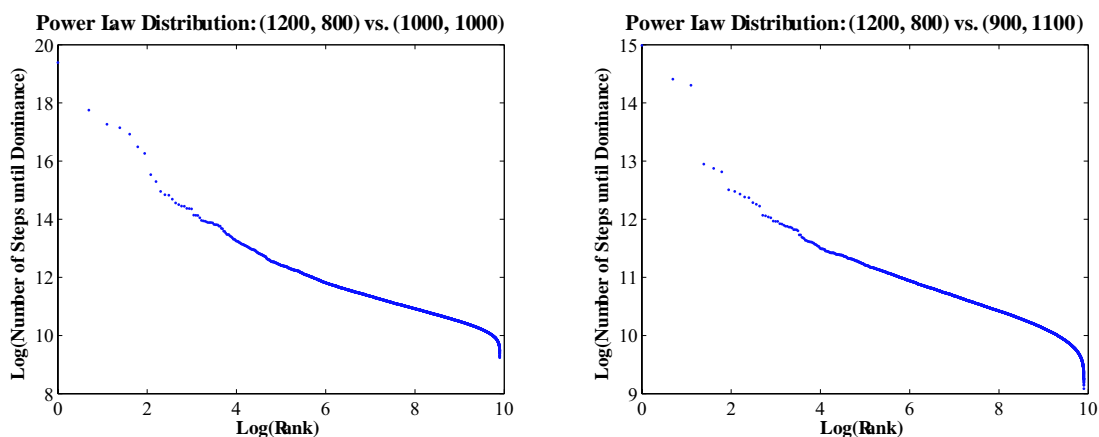

Figure 3: We use log-log plots to illustrate the power law distribution of time until dominance. Here, $p = q = 2$ and $\alpha = 2/3$. In the first diagram, we start with $(1200, 800)$ and $(1000, 1000)$ for network A and B. In the second diagram, we start with $(1200, 800)$ and $(1100, 900)$ for network A and B.

5 Simulation Results

5.1 Three Cases

We have shown that the value of pq, with a phase transition at $pq = 1$, determines the market structure of two-sided markets in the large-time limit. We demonstrate this behavior in simulations of the model, where we sequentially simulate the random decisions of new arrivals. Figure 4 illustrates simulation results for different values of pq. For Figure 4, we use the same initial states, $(120, 120)$ for network A and $(80, 80)$ for network B, and the same parameter $\alpha = 2/3$ that determines the probability a new customer is a buyer. Network A has advantages on both buyer and seller sides initially. We present the results for different values of pq after $300,000$ new arrivals. In each diagram in Figure 4, we run the same setup 500 times, and use the market shares of network A on two sides as the (x, y) coordinates. As expected, even with the same setup, results vary from one simulation run to another because of the randomness in our probabilistic model. For instance, in the case where $pq > 1$ the market share of network A on the seller side varies from 70% to near 100% in different simulation runs. These simulation results are consistent with our theoretical results. Since network A has advantages on both the buyer and seller sides, when $pq > 1$ the advantages increase over time. For $pq = 1$, the equilibrium market shares fall in a region close to the initial values. For $pq < 1$, the advantages of network A vanish. Network A and network B have almost equal market shares on both buyer and seller sides.

Figure 5 illustrates typical convergence paths for different values of pq. Note that the market shares tend to fluctuate widely initially and then become stable. This is consistent with the intuition that as install bases of both sides become larger, a new arrival will have a smaller impact on market shares. Similarly, even though theoretically the market shares of network A is expected to converge to 1 when $pq > 1$, and to converge to 50% when $pq < 1$, as the sizes of the installed bases increase, we need to have an increasing amount of new arrivals in order to change the market shares by the same amount. Finally, in the middle case where $pq = 1$, a simple calculation reveals that the market share appears to be converging to a point where $r_b^q = r_s$, as predicted.
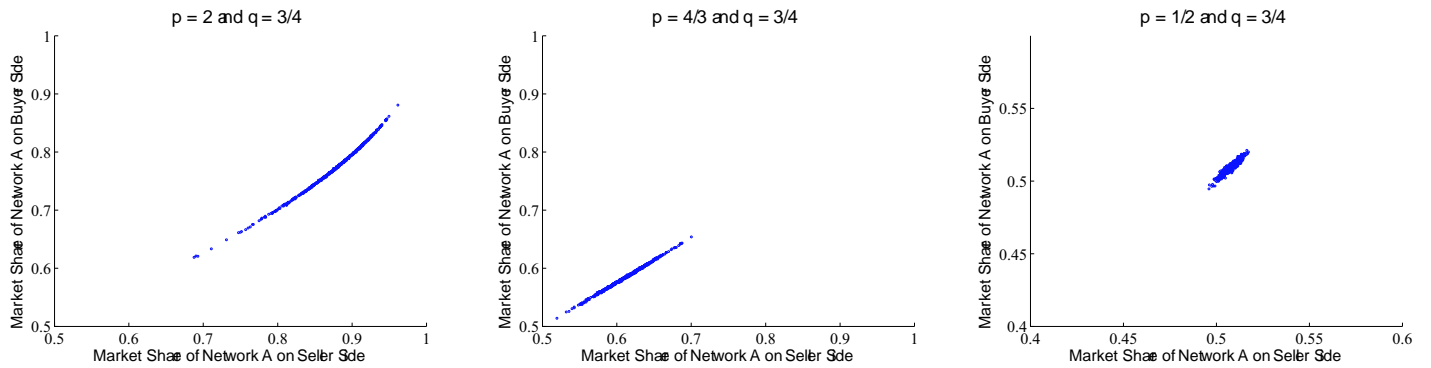
Figure 4: Simulation results for $pq > 1$, $pq = 1$ and $pq < 1$ with the same initial states $(120, 120)$ and $(80, 80)$ and $\alpha = 2/3$. In each diagram, we run the same setup 500 times, and use the market shares of network A on two sides after $300,000$ new adopters as the (x, y) coordinates. We use $(p = 2, q = 3/4)$, $(p = 4/3, q = 3/4)$, and $(p = 1/2, q = 3/4)$ as the p, q values for each of these three diagrams.
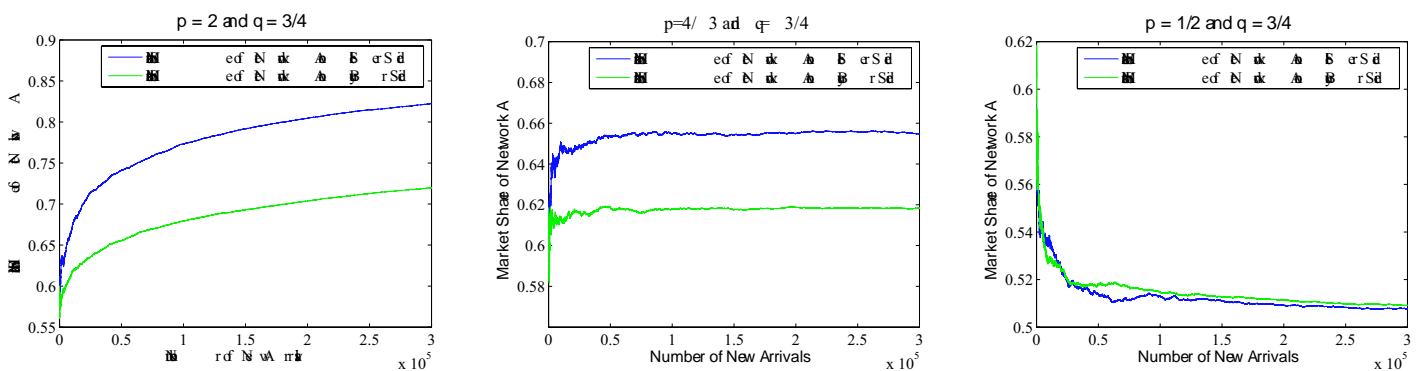

Figure 5: Typical convergence paths for different values of pq. We use the same simulation setup as in Figure 4 for each diagram. We illustrate for different pq values how the market shares of network A on both sides change as new agents continue to arrive.

14

5.2 The Monopoly Case

In Figure 6, we present the cumulative distribution for the number of buyers in network A after $1,500,000$ new arrivals. In the first two diagrams, we start with one agent in each side of the networks. Since $\alpha = 2/3$ here, the total number of buyers is about $1,000,000$. In the first diagram, we use $p = q = 1.1$. While there is obvious tendency towards monopoly, there is still a reasonable probability that one network will not overwhelm the other. For example, the probability that one network has over 80% of buyers is less than 80%. In contrast, in the second diagram where $p = q = 1.5$, one network has almost all of the new buyers. The probability that one network receives less than 200 agents on both sides is 68.9%. This concentration results from the dramatic effect inequality creates at the beginning of the process. When p and q are large and the total number of agents in the market is small, even small leadership implies a huge advantage. In the third diagram, we begin with 100 agents in each side of the networks and let $p = q = 1.5$. We see the curve appears more similar to the $p = q = 1.1$ case. These results illustrate that small markets and large p and q indicate quick time to dominance, but otherwise time to dominance can be very slow, and the resulting market distribution can be highly variable.
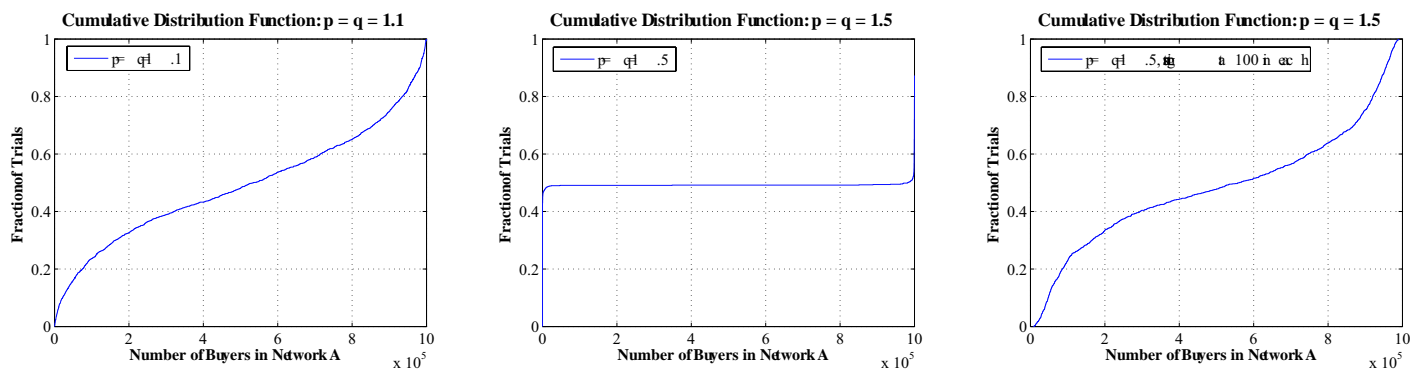

Figure 6: Cumulative distribution functions of the number of buyers in network A after $1,500,000$ new arrivals, $\alpha = 2/3$. We start with $(1,1)$ and $(1,1)$ for network A and B in the first two diagrams, and use $p = q = 1.1$ and $p = q = 1.5$ respectively. In the third diagram, we start with $(100,100)$ and $(100,100)$ for network A and B, and use $p = q = 1.5$.

5.3 Relative Arrival Rates

We illustrate the effect of different relative arrival rates between buyers and sellers in the case of $pq > 1$ in Figure 7. Recall that in our framework the parameter α is used to determine the probability a new entrant is a buyer. When a network has an advantage on both sides, different arrival rates do not affect its ultimate chance of dominance significantly. Therefore, here we choose the initial states to be $(120, 80)$ for network A and $(100, 100)$ and for network B. That is, network A has more than 50% market share on the buyer side but has less than 50% market share on the seller side. In addition, we assume $p = q = 2$. We allow α to increase from 0.1 to 0.9 by an increment of 0.05. For simulation purposes, we say here that a network achieves dominance when it has more than 60% market shares on both sides. For each value of α, we run the same simulation $1,000$ times. The probability of becoming the dominant network is computed by the number of times the network achieves dominance divided by $1,000$. Figure 7 indicates that the probability that network A achieves dominance decreases as α increases. The result is intuitively clear. When α is small, i.e., sellers arrive more frequently than buyers, network A can catch up with network B on the seller side while still maintaining its advantage on the buyer side. When α is large, network A loses its advantage on the buyer side before it catches up with network B on the seller side.

6 Implications

Bayus and Shankar (2003)'s study on the video game industry shows that the strength of the indirect network effects in addition to the size of the networks play important role in the success of the game consoles. Our results are consistent with their finding. More importantly, we show that the ultimate market structure depends significantly on the underlying strength of the indirect network effects. Knowing the value of pq is therefore of strategic importance for intermediaries. If $pq > 1$, it is very important for an intermediary to gain and maintain the leadership position, as the market will eventually tend to a monopoly. If $pq < 1$, an intermediary can still survive and even thrive in the long term even if it does not have the largest market share on both sides.

The parameters p, q, and α can be estimated from market data using standard techniques, giving insight into future market behavior. The parameter α can tracked by ex-
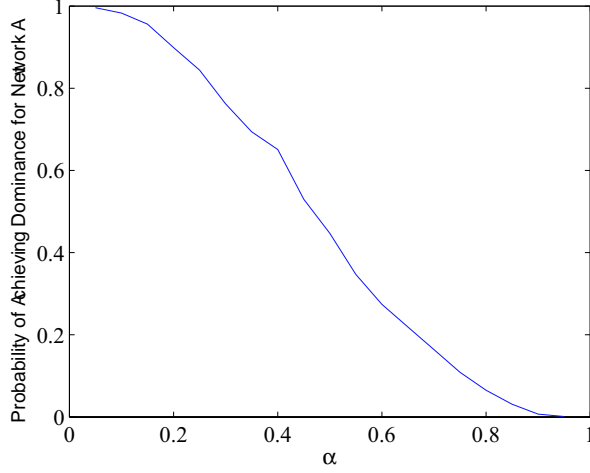
Figure 7: The effect of relative arrival rates of buyers and sellers on the probability of achieving dominance. Here the initial states are $(120, 80)$ and $(100, 100)$ for network A and B. We also let $p = q = 2$. For simulation purpose, we consider a network achieves dominance when it has more than 60% market shares on both sides. For each value of α, we run the same simulation $1,000$ times and compute the probability of achieving dominance for network A.

amining the rates of incoming buyers and sellers. The parameter p can be estimated by considering the fraction of buyers that join network A over a small period of time and solving $r = \frac{s_A^p}{s_A^p + s_B^p}$: here $p = \frac{\ln r - \ln(1-r)}{\ln s_A - \ln s_B}$. (Similar analysis can be done to estimate q.) While s_A and s_B may vary over time, when s_A and s_B are sufficiently large compared to the sampling period, the changes in s_A and s_B will be sufficiently small that they can be treated as constant for estimation purposes. (More complex techniques can be used when this is not the case.)

Alternatively the behavior of the stochastic process can be closely approximated by log-linear demand functions similar to those used in Keilbach and Posch (1998). The demand function for network k at time t at the buyer side is $D_k^b(t) = C \cdot (ms_k^s(t))^p$ and at the seller side is $D_k^s(t) = C' \cdot (ms_k^b(t))^q$, where $ms_k^b(t)$ and $ms_k^s(t)$ are the market shares of network k at the buyer and seller sides at time t respectively. C and C' here are constants used for scaling purpose and their ratio represents the relative arrival rates between the two sides. The parameters p and q, and the relative arrival rates between the two sides can then be estimated in a regression framework. An intermediary may use such estimates along with the current state of each network in the market to gain insights into the expected evolution path of its network.

17

In addition, an intermediary may also want to use strategies such as pricing, advertising and product differentiation to alter the evolution of the market share. Although our baseline model does not incorporate these strategic variables, it can be used to examine the possible effects of these actions. Consider a scenario in which two intermediary technologies are of the same quality initially, and at time t the intermediary in network A improves the quality of its technology. Suppose the quality improvement leads to a factor of Q ($Q > 1$) increase in the utility network A provides to its two sides.

We may re-derive the equations for dr_s/dt and dr_b/dt to obtain an appropriate state diagram. An example is shown in Figure 8 for a case where $pq > 1$. The dashed curves represent the boundaries between differing trajectory behaviors for the state diagram before the quality improvement and the solid curves the boundaries after the quality improvement. The quality factor changes the relevant boundary equations to $Qr_s^p = r_b$ and $Qr_b^q = r_s$. It is worth noting that after this change the state $(1, 1)$, where the numbers of sellers and buyers are the same for both networks, is no longer a fixed point; instead the heuristic predicts that network A will dominate, as intuition would suggest.

Using this type of modeling and analysis, an intermediary can determine the possible effects of strategies such as pricing, advertising, and product differentiation on future market share. In particular, an intermediary might determine what strategies would allow it to achieve a trajectory leading to market dominance.

7 Conclusions and Future Research

We have developed a novel dynamic model to show how the power of positive indirect network effects can shape the market structure in the setting of two-sided markets. Our results indicate that the ultimate market structure depends on the product of the market share elasticities of the two sides, regardless of initial states and the relative arrival rate of buyers and sellers. When the product is greater than one, one network will eventually become a monopoly. When the product is less than one, the market tends to a state where sellers and buyers are equal for both networks. In the monopoly case, we have illustrated that the time until dominance follows a power law distribution. In addition, when the overall strength of the indirect network effects is not strong enough, there is a high probability that one network will not overwhelm the other. Intermediaries in two-sided
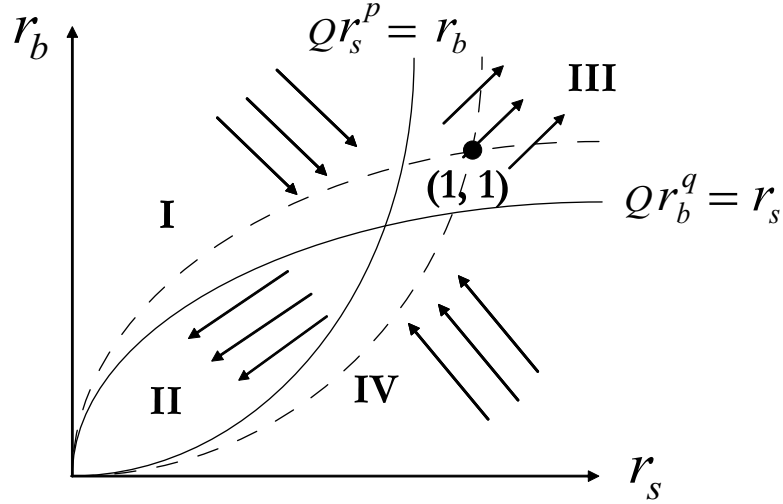
18

Figure 8: A state diagram showing the effect of a proposed quality improvement when $p > 1$ and $q > 1$. The dashed curves represent the boundaries between trajectory types before the quality improvement and the solid curves represent the boundaries after the quality improvement. Arrows are used to indicate the directions of trajectories for states in each region after the quality improvement, as determined by the signs of $\frac{dr_s}{dt}$ and $\frac{dr_b}{dt}$.

markets can then use these results to make more informed decisions after estimating the strengths of indirect network effects.

There are several issues that remain to be investigated. In the state diagrams in Figure 2, we are able to trace a trajectory for every state as t increases. These trajectories allow us to predict long-term evolution of a solution for System A. In the case of $pq > 1$, for states where one network has an advantage in buyers and the other network has an advantage in sellers, given a specific set of initial conditions one can use the equations of System A to predict which network will eventually dominate. We would like to have an easily expressible formula or conditions, based on p, q, α, and the initial state, that determines which network will reach monopoly in System A.

Second, our results on time to dominance have been based on simulation runs or running the differential equations. Ideally, we would similarly like to derive a formula giving the order of magnitude or the exact expected time to dominance based on p, q, α, and initial state.

Third, in our simulation runs, we noticed that given the size of a network and $pq > 1$, there appears to be an optimal state for the network, in which the probability of achieving dominance is maximized. The optimal state does not depend on the state of the other

19

competing network. We do not have a ready explanation for this observation. Clearly, this result, if true, has important strategic implications for intermediaries, and should be explored.

Most of these questions would be much easier to answer if we could design stronger analysis techniques for this problem. As we have discussed, the exponential embedding technique for the simpler balls-and-bins setting (corresponding to one-sided markets) allows very powerful results to be obtained for similar questions. Unfortunately, we have not found a method of generalizing this technique to our model of two-sided markets. We have been able to generalize weaker discrete analysis of Drinea et al. (2002), as shown in the appendix, to obtain rigorous results, but we believer stronger results are possible. Finding an appropriate generalization of the exponential embedding approach to our model is a key step for future research. Such a generalization would likely also be useful for developing further generalizations and variations of these types of models for network effects. Alternatively, it may be worth considering whether there are changes we could make to the model that would make it more amenable to this analysis technique, while maintaining its spirit.

Appendix: Combinatorial Proofs of Increasing Returns When $p > 1$ and $q > 1$

Recall that in the balls-and-bins setting, there are no buyers and sellers; instead, balls join the network one at a time, so that at each time step a ball joins bin i with probability $\frac{x_i^p}{\sum_{j=1}^n x_j^p}$, where x_i is the number of existing balls in bin i. In this setting, the exponential embedding technique yields very clean and natural rigorous proofs of system behaviors, including that when $p > 1$ a monopoly occurs, in the very strict sense that after a certain point only one bin will receive any further balls (Davis 1990, Mitzenmacher et al. 2004, Oliveira 2004). Unfortunately, we have not found that this approach extends to two-sided markets. Using the weaker methods based of Drinea et al. (2002) based on large deviations, we can give rigorous combinatorial results for the case $p > 1$ and $q > 1$. These techniques could be used to examine other parameter values, and a more refined analysis could be done by for example embedding the process into a gambler's ruin type Markov chain. Finding simpler and stronger methods, based on the exponential embedding or other techniques, that give rigorous and precise results remains an important open prob-

lem.

We sketch the approach; see Drinea et al. (2002) for more details. Suppose that at time t

$$b_A(t) \geq \left(\frac{1}{2} + \epsilon\right) \alpha t, \qquad\qquad b_B(t) \leq \left(\frac{1}{2} - \epsilon\right) \alpha t,$$

$$s_A(t) \geq \left(\frac{1}{2} + \epsilon\right) (1 - \alpha)t, \qquad s_B(t) \leq \left(\frac{1}{2} - \epsilon\right) (1 - \alpha)t.$$

for some $\epsilon < 1/4$. That is, suppose that A leads in both buyers and sellers as above. Then if $p > 1$ and $q > 1$, we can show that with high probability these leads will increase steadily over time, in the sense that if t is sufficiently large, then with high probability there exists a constant $c > 1$ such that

$$b_A(2t) \geq \left(\frac{1}{2} + c\epsilon\right) \alpha t, \qquad\qquad b_B(2t) \leq \left(\frac{1}{2} - c\epsilon\right) \alpha t,$$

$$s_A(2t) \geq \left(\frac{1}{2} + c\epsilon\right) (1 - \alpha)t, \qquad s_B(2t) \leq \left(\frac{1}{2} - c\epsilon\right) (1 - \alpha)t.$$

That is, after t more steps the "gap" between the two competitors will in some sense increase by a constant factor. It is worth noting that we must be careful in stating this result. For example, if $s_A(t)$ is much bigger than $s_B(t)$, and $b_A(t)$ is very close to $b_B(t)$, then because the incoming seller's choice of network depends on the buyer distribution, the ratio between s_A and s_B may come closer together. We cannot say that A will necessarily always increase the lead in both buyers and sellers simultaneously, and it is for this reason we use the same ϵ in the above equations.

We consider what happens after γt additional agents (buyers or sellers) arrive, for some suitably small constant γ. Over these γt steps, for $t \leq u \leq t + \gamma t$, we have that

$$\frac{b_A(u)}{b_B(u)} \geq \frac{1/2 + \epsilon}{1/2 - \epsilon + \gamma} = 1 + \frac{4\epsilon - 2\gamma}{1 - 2\epsilon + 2\gamma}.$$

Hence for $q > 1$ and $\gamma < \epsilon$

$$\left(\frac{b_A(u)}{b_B(u)}\right)^q \geq \left(1 + \frac{4\epsilon - 2\gamma}{1 - 2\epsilon + 2\gamma}\right)^q > 1 + \frac{4q\epsilon - 2\gamma}{1 - 2\epsilon + 2\gamma}.$$

Let us use the notation

$$R_v = 1 + \frac{4q\epsilon - 2\gamma}{1 - 2\epsilon + 2\gamma}$$

and

$$R_t = \frac{1/2 + \epsilon}{1/2 - \epsilon}.$$

Now for suitably small ϵ and γ we have for any $t \leq u \leq t + \gamma t$ that

$$\left(\frac{b_A(u)}{b_B(u)}\right)^q > R_v > R_t.$$

For example, we the above holds for any $\epsilon < 1/4$ and any $\gamma < 2\epsilon(q-1)/3$.

It follows that the ratio of the probabilities that an incoming seller chooses network A and that an incoming seller chooses network B is strictly greater than R_t over the course of the γt steps. In particular, the probability that an incoming seller chooses network A is at least $1/2 + c_1\epsilon$ for some constant $c_1 > 1$. For any constants δ_1 and δ_2, if t sufficiently large, then using standard Chernoff bounds, we have that at least $1 - \alpha - \delta_1$ new arrivals are sellers with high probability, at that of these sellers, at least a $1/2 + c_1\epsilon - \delta_2$ fraction of them join network A. Hence the increase in s_A over these γt steps is, with high probability, at least

$$\left(\frac{1}{2} + c_1\epsilon - \delta_2\right)(1 - \alpha - \delta_1)\gamma t \geq \left(\frac{1}{2} + c_2\epsilon\right)(1 - \alpha)\gamma t$$

for some constant $c_2 > 1$ (where δ_1 and δ_2 have been chosen to be suitably small).

We may repeat this argument for $1/\gamma$ intervals each of γt arriving agents, to find that over the next t steps, the number of new sellers arriving at network A is at least

$$\left(\frac{1}{2} + c_2\epsilon\right)(1 - \alpha)t$$

with high probability. This then gives that

$$s_A(2t) \geq \left(\frac{1}{2} + c\epsilon\right)(1 - \alpha)t$$

with $c = (1 + c_2)/2$. An entirely symmetric argument holds to show that

$$b_A(2t) \geq \left(\frac{1}{2} + c\epsilon\right)\alpha t$$

for some (possibly different) constant c as well; taking the minimum of these two constants gives the final result.

Finally, a similar approach can be used to prove a similar statement as the concentrations approach 1, as in Drinea et al. (2002) for more details. That is, suppose that at time t,

$$b_A(t) \geq (1 - \epsilon)\, \alpha t, \qquad\qquad b_B(t) \leq \epsilon \alpha t,$$

$$s_A(t) \geq (1 - \epsilon)\, (1 - \alpha) t, \qquad\qquad s_B(t) \leq \epsilon(1 - \alpha) t.$$

for some suitably small ϵ. That is, suppose that A leads dramatically in buyers and sellers, having all but an ϵ fraction of both. Then if $p > 1$ and $q > 1$, we can show that with high probability there exists a constant $c > 1$ such that

$$b_A(2t) \geq \left(1 - \frac{\epsilon}{c}\right) \alpha t, \qquad\qquad b_B(2t) \leq \frac{\epsilon}{c} \alpha t,$$

$$s_A(2t) \geq \left(1 - \frac{\epsilon}{c}\right) (1 - \alpha) t, \qquad\qquad s_B(2t) \leq \frac{\epsilon}{c}(1 - \alpha) t.$$

That is, the all-but-ϵ advantage improves to an all-but-ϵ/c advantage.

References

Armstrong, Mark. forthcoming. Competition in two-sided markets. *Rand Journal of Economics* .

Arthur, Brian W. 1989. Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal* **99** 116–131.

Arthur, Brian W. 1994. *Increasing Returns and Path Dependence*. The University of Michigan Press.

Arthur, Brian W., Yuri M. Ermoliev, Yuri M. Kaniovski. 1983. On generalized urn schemes of the pólya kind. *Translated from the Russian in Cybernetics* **19** 61–71.

Auriol, Emmanuelle, Michel Benaïm. 2000. Standardization in decentralized economics. *American Economic Review* **90**(3) 550–570.

Bayus, Barry L., Venkatesh Shankar. 2003. Network effects and competition: An analysis of the home video game industry. *Strategic Management Journal* **24** 375–384.

Benaïm, Michel, Jörgen W. Weibull. 2003. Deterministic approximation of stochastic evolution in games. *Econometrica* **71**(3) 873–903.

Caillaud, Bernard, Bruno Jullien. 2003. Chicken & egg: Competition among intermediation service providers. *RAND Journal of Economics* **24** 309–328.

Chakravorti, Sujit, Roberto Roson. 2004. Platform competition in two-sided markets: The case of payment networks. *Federal Reserve Bank of Chicago Emerging Payments Occasional Paper Series* **2004-09**.

Chou, Chien-fu, Oz Shy. 1993. Partial compatibility and supporting services. *Economics Letters* **41** 193–197.

Church, Jeffrey, Neil Gandal. 1992. Network effects, software provision and standardization. *Journal of Industrial Economics* **40**(1) 85–104.

Clements, Matthew T., Hiroshi Ohashi. 2005. Indirect network effects and the product cycle: Video games in the U.S., 1994-2002. *Journal of Industrial Economics* **LIII**(4) 515–542.

Davis, Burgess. 1990. Reinforced random walks. *Probability Theory and Related Fields* **84** 203–229.

Drinea, Eleni, Alan Frieze, Michael Mitzenmacher. 2002. Balls in bins processes with feedback. *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms* 308–315.

Evans, David S. 2003. Some empirical aspects of multi-sided platform industries. *The Review of Network Economics* **2**(3) 191–209.

Evans, David S., Andrei Hagiu, Richard Schmalensee. 2004. A survey of the economic role of software platforms in computer—based industries. Discussion papers, Research Institute of Economy, Trade and Industry (RIETI). Available at http://ideas.repec.org/p/eti/dpaper/04032.html.

Farrell, Joseph, Garth Saloner. 1985. Standardization, compatibility, and innovation. *RAND Journal of Economics* **16** 70–83.

Farrell, Joseph, Garth Saloner. 1986. Installed base and compatibility: Innovation, product preannouncements, and predation. *American Economic Review* **76** 940–955.

Fudenberg, Drew, David K. Levine. 1998. *The Theory of Learning in Games*. The MIT Press, Cambridge, MA.

Gabszewicz, Jean J., Xavier Y. Wauthy. 2004. Two-sided markets and price competition with multi-homing. *Mimeo, CORE, Louvain-la-Neuve University* .

Guthrie, Graeme, Julian Wright. 2003. Competing payment schemes. *Working Paper No. 0311, Department of Economics, National University of Singapore* .

Kandori, Michihiro, George J. Mailath, Rafael Rob. 1993. Learning, mutation, and long run equilibria in games. *Econometrica* **61**(1) 29–56.

Katz, Michael L., Carl Shapiro. 1985. Network externalities, competition, and compatibility. *American Economic Review* **75**(3) 424–440.

Keilbach, Max, Martin Posch. 1998. Network externalities and the dynamics of markets. *Working Paper, International Institute for Applied Systems Analysis* .

Khanin, Kostya, Raya Khanin. 2001. A probabilistic model for the establishment of neuron polarity. *Journal of Mathematical Biology* **42**(1) 26–40.

Kurtz, Thomas G. 1981. *Approximation of Population Processes*. SIAM.

Mitzenmacher, Michael. 1996. *The Power of Two Choices in Randomized Load Balancing*. Ph. D. Thesis, U.C. Berkeley.

Mitzenmacher, Michael, Roberto Oliveira, Joel Spencer. 2004. A scaling result for explosive processes. *Electronic Journal of Combinatorics* **11**(1) R31.

Nair, Harikesh, Pradeep Chintagunta, Jean-Pierre Dubé. 2004. Empirical analysis of indirect network effects in the market for personal digital assistants. *Quantitative Marketing and Economics* (2) 23–58.

Oliveira, Roberto. 2004. Preferential attachment. *Ph.D. Dissertation, Department of Mathematics, New York University* .

Rochet, Jean-Charles, Jean Tirole. 2003. Platform competition in two-sided markets. *Journal of European Economic Association* **1** 990–1029.

Weibull, Jörgen W. 1995. *Evolutionary Game Theory*. The MIT Press, Cambridge, MA.

Wormald, Nick C. 1995. Differential equations for random processes and random graphs. *Annals of Applied Probability* **5** 1217–1235.