

Social Influence and Evolution of Market Share

Simla Ceyhan, Mohammad Mousavi, and Amin Saberi

Abstract. We propose a model for the evolution of market share in the presence of social influence. We study a simple market in which the individuals arrive sequentially and choose one of a number of available products. Their choice of product is a stochastic function of the inherent quality of the product and its market share.

Using techniques from stochastic approximation theory, we show that market shares converge to an equilibrium. We also derive the market shares at equilibrium in terms of the level of social influence and the inherent quality of the products.

In a special case, in which the choice model is a multinomial logit model, we show that inequality in the market increases with social influence and that with strong enough social influence, monopoly occurs. These results support the observations made in the experimental study of cultural markets in [Salganik et al. 06].

I. Introduction

It has been observed consistently in various social and economic settings that individuals generally look to other people when making decisions [Cialdini 01]. This phenomenon is even more pronounced in online social settings such as YouTube and Digg, in which clips or news items that have a high “market share,” or high number of views, are chosen more frequently [Cha et al. 07, Ost et al. 08].

In this paper, we study the effect of such social influence on the evolution of market share. We consider a model in which individuals sequentially choose one of a number of available products based on its inherent quality as well as its market share. The market share of a product is defined as the fraction of people who have chosen that product before. Also, the decision of every individual is a stochastic function of these two parameters.

This setting gives rise to a number of interesting questions. Do these markets converge to an equilibrium? Is the outcome (or the equilibrium) of this market predictable? Can we observe a significant difference in the outcome if we increase the level of social influence? What is the effect of the inherent appeal, or quality, of a product on its market share?

Our main result shows that market shares converge to an equilibrium if there is a sufficiently large number of participants. Furthermore, it is possible to derive the equilibrium as a solution of an ordinary differential equation (ODE). The proof of this theorem uses techniques from stochastic approximation theory that relate the limit behavior of a stochastic process to the limit behavior of a differential equation.

We also study the rate of convergence to the equilibrium. We prove that the difference between the stochastic process and its corresponding ODE converges weakly to a Gaussian diffusion. More specifically, for a sufficiently large number of people in the market, with high probability, the difference between the market share and the equilibrium point is less than c/\sqrt{n} , where c depends on the social influence. Furthermore, for some special cases of our model, we observe a change in equilibrium with the level of social influence.

Our results are applicable to a very general class of functions. In fact, they imply some of the classic results on balls-and-bins processes studied in probability theory [Oliveira 08, Chung et al. 03] as a special case. In balls-and-bins processes, balls are sequentially thrown into bins in such a way that the probability that a bin with n balls receives the next ball is proportional to $f(n)$, for some feedback function f of type $f(n) = n^p$ for $p > 0$. It has been shown that when $p > 1$, almost surely there is some bin that gets all but finitely many balls. This is what we refer to as the *monopoly* case. When $p < 1$, the asymptotic fraction of balls in each bin is the same, whereas $p = 1$ becomes the classic Pólya urn model. Our analysis implies these results for $p < 1$ and $p > 1$.

Another special case of our model is the multinomial logit model, which is widely used for modeling social influence in the economics literature. There are also many empirical studies that use the logit model to analyze social influence [Sorensen 06, Glaeser et al. 96, Baddeley 07, Dubin 95, Paez et al. 08]. In this special case, we can measure the effect of social influence on the market share more directly. For example, we can show that the number of equilibria increases

with the social influence. This can be interpreted as an increase in the unpredictability of the market.

In this special case, we also study the effect of social influence on inequality. We observe that in the case of weak social influence, the quality of the products has a significant effect on the final market shares. However, with strong social influence, monopoly occurs; that is, one of the products, which does not necessarily have the highest quality, gets the largest market share, while for the rest, the market share becomes almost zero. Both of these results are consistent with the observations made in the experimental work [Salganik et al. 06].

1.1. Related Literature

The effect of social influence on collective behavior has been studied extensively within the fields of marketing, sociology, and economics [Chandukala et al. 08, Wierenga 08, Schelling 78, Follmer 74, Grossman and Stiglitz 76, Young 98]. More recently, social learning models have been incorporated into standard models of economic decision-making by a growing number of theoretical studies.

An extensive literature in economic theory uses rational or Bayesian learning as the underlying hypothesis for explaining herding or social learning. In this setting, individuals are assumed to have imperfect information about the quality of their choices. Therefore, their decision is based on their imperfect information as well as on what they infer from previous decisions made by others [Banerjee 92, Bikhchandani et al. 92, Ellison and Fudenberg 93, Chamley 03, Acemoglu et al. 08, Acemoglu et al. 10].

There are also economic models that study social influence as the dynamics of a coordination game. There, the hypothesis is that an individual will have a higher utility for taking the same action as those who came before him. The interested reader can see [Young 98, Montanari and Saberi 08] for more information.

The model proposed in this paper is in the following reduced form: instead of modeling the information or the incentive structure of every individual, we assume that his or her decision is a monotone function of the perceived quality of the option as well as its market share. We expect most game-theoretic models described above to have such a monotonicity.

Despite many theoretical studies, social effects have been hard to quantify precisely in practice. This is mostly due to the difficulty in drawing inferences from the data [Manski 00]. One can still find empirical studies on social influence in many different contexts such as crime [Glaeser et al. 96], labor supply [Woittiez and Kapteyn 98], stock market participation [Hong et al. 04], and choice of health plans [Bertrand et al. 00].

Our work is mostly related to the experimental study [Salganik et al. 06]. The authors study a web-based music market with over 14,000 consumers to understand the effect of social influence in cultural markets. The main conclusion of the paper is that inequality and unpredictability in the market increase with social influence. Also, recently, we have come across a study of an online music community called *The Hype Machine*. Similarly, it is shown in [Ramaprasad and Dewan 09] that being able to observe others' adoption decisions positively influences subsequent consumption decisions, and this is consistent across independently released and major-label-released music.

The organization of the rest of the paper is as follows. In Section 2, we present our model and analytical results. In Section 3, we focus on the logit choice model and look at the change in the equilibrium with social influence. Our conclusions appear in Section 4. There follow three appendices, Sections 5–7, containing statements and proofs of several needed lemmas and theorems.

2. The Model

We consider a market with m products. At each time step n , a new participant enters the market. After observing the market share of each product, he chooses a product $i \in \{1, \dots, m\}$ that satisfies

$$i \in \operatorname{argmax}_j (Jh(\phi_j(n)) + f_j + \epsilon_j).$$

Here J is a parameter that measures the social influence, $\phi_j(n)$ is the market share of the product j at time n , h is a monotone smooth function, f_j is the inherent quality of the product j , and ϵ_j is the noise. We call the sequence $\phi(n)$ of market shares generated according to this model the *market share sequence*. According to our model, each customer's decision is influenced by what others have chosen previously (the market share), the social influence (for $J > 0$), and the quality of the products. Once the stochastic terms are realized, each customer chooses an alternative that maximizes the above equation. From now on we will use vectors in our notation: $\phi(n) = (\phi_1(n), \dots, \phi_m(n))$ is the market share vector at time n , and $f = (f_1, \dots, f_m)$ is the quality vector.

The above model can be considered an instance of the *additive random utility model* in discrete choice theory, which is frequently used by economists [McFadden 80, Ben-Akiva and Lerman 85, Anderson et al. 92]. In this model, an agent must choose from a set of alternatives $A = \{1, \dots, m\}$ offering some base utility, in our case

$$x_j = Jh(\phi_j(n)) + f_j,$$

and some stochastic utility ϵ_j . The noise vector $\epsilon = (\epsilon_1, \dots, \epsilon_m)$ has strictly positive density everywhere in \mathbb{R}^m . Once the stochastic terms are realized, each agent chooses the alternative whose total utility is maximal.

Let $X = (x_1, x_2, \dots, x_m)$. We define the *choice probability function* as

$$L : \mathbb{R}^m \rightarrow \mathbb{R}^m$$

$$L_i(X) = P \left(\operatorname{argmax}_j x_j + \epsilon_j = i \right),$$

where $L_i(X)$ gives the probability that the agent chooses alternative i .

Hence, in our model the $(n + 1)$ th customer will choose product i with probability

$$C_i^J(\phi(n)) = L_i(Jh(\phi(n)) + f) = P \left(\operatorname{argmax}_j Jh(\phi_j(n)) + f_j + \epsilon_j = i \right). \quad (2.1)$$

The most common choice probability function used to model herding is the *logit choice function* [Vulcano et al. 10, Ratliff et al. 08, Sorensen 06, Glaeser et al. 96]. If ϵ_j are independent identically distributed random variables with the extreme value distribution such that

$$F(\epsilon) = \exp(-\exp(-\eta^{-1}\epsilon - \gamma)),$$

where γ is Euler's constant, then the choice probability function is the logit choice function

$$L_i(X) = \frac{e^{\eta^{-1}x_i}}{\sum_j e^{\eta^{-1}x_j}}.$$

A major advantage of the logit model is that it is analytically tractable. In particular, it has a closed-form expression for the choice probabilities. As a side note, let us give two examples of our model when ϵ_j are independent and identically distributed random variables with the extreme value distribution:

Example 2.1. Let h be the identity function, and $\eta = 1$. Then

$$C_i^J(\phi(n)) = \frac{e^{J\phi_i(n) + f_i}}{\sum_j e^{J\phi_j(n) + f_j}}.$$

Example 2.2. Let $h = \ln(x)$ and $\alpha_i = \exp(f_i)$. Then our model becomes a generalization of the preferential attachment

$$C_i^J(\phi(n)) = \frac{\alpha_i(\phi_i(n))^J}{\sum_j \alpha_j(\phi_j(n))^J}.$$

For $J = 1$, the above model is used in [Borgs et al. 07] as preferential attachment with fitness. Also, if the fitness of each product is the same, our model becomes the balls-in-bins process with feedback, where the feedback function is $f(x) = x^J$.

For the rest of the paper, we will refer to our model as $C(\phi(n))$ to simplify the notation.

2.1. Main Theorems

Using the above model, we study the behavior of market shares when a sufficiently large number of customers have visited the market. In the first theorem, we will show that in the general form, the stochastic process of the market share asymptotically follows a deterministic path, which is the solution of an ODE. Next, we will show that the market share converges to an equilibrium at infinity in the special case that h is linear. In the third theorem, we study the rate of convergence of the equilibrium.

Theorem 2.3. *Let $C : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be the choice probability function defined in (2.1), where $\phi(n)$ is the market share at time n . Then there exists a sequence $(\bar{t}_k)_{k \geq 1}$ such that as $\bar{t}_k \rightarrow \infty$,*

$$\lim_{k \rightarrow \infty} |\phi(k) - \bar{\phi}(\bar{t}_k)| = 0, \quad (2.2)$$

where $\bar{\phi}$ is the solution of the following ODE:

$$\frac{d}{dt} \bar{\phi}(t) = C(\bar{\phi}(t)) - \bar{\phi}(t). \quad (2.3)$$

Theorem 2.3 shows that in the limit, the evolution of market shares is governed by the ODE in (2.3). The theorem has a simple interpretation. It shows that the steady state of market shares can be derived from a deterministic and continuous-time model. In this model, at every time t , a continuum of customers comes to the market and chooses a portion of each product i as a function $C(\phi_i(t))$ of its market share. It is straightforward to observe that the rate of change of market share in this new model can be derived from (2.3). Theorem 2.3 shows that the original model asymptotically behaves like this continuous-time model.

In the proof, we first construct a continuous-time interpolation of the market share sequence. Then, we apply some results from stochastic approximation theory to show that this function asymptotically behaves like an ODE.

Proof. For each customer n , define the indicator $I^n \in \mathbb{R}^m$, representing the product selected by customer n . In other words, $I_i^n = 1$ if the customer selects product i , and $I_i^n = 0$ otherwise.

Suppose product i has been chosen N_i^n times by the first n customers. Let N^n be a vector with m components such that component i is N_i^n . Here we assume that each customer buys exactly one product. Therefore,

$$N^{n+1} = N^n + I^{n+1}.$$

By definition, market share is $\phi(n) = \frac{1}{n}N^n$, and $C(\phi(n)) = E[I^{n+1}|\phi(n)]$. Hence,

$$\begin{aligned} \phi(n+1) &= \frac{n}{n+1}\phi(n) + \frac{1}{n+1}I^{n+1}, \\ E[\phi(n+1) - \phi(n)|\phi(n)] &= -\frac{1}{n+1}\phi(n) + \frac{1}{n+1}E[I^{n+1}|\phi(n)] \\ &= -\frac{1}{n+1}\phi(n) + \frac{1}{n+1}C(\phi(n)) \\ &= \frac{1}{n+1}(C(\phi(n)) - \phi(n)). \end{aligned} \quad (2.4)$$

Let $a_n = \frac{1}{n+1}$, $g(x) = C(x) - x$, and $d_n = I^{n+1} - C(\phi(n))$. Now we can rewrite $\phi(n+1)$ as

$$\phi(n+1) = \phi(n) + a_n g(\phi(n)) + a_n d_n. \quad (2.5)$$

Next, we will use the approach of the theory of stochastic approximations from [Kushner and Yin 03]. Let us start with showing that the assumptions necessary to use ODE methods are satisfied. From (2.4), $E[a_n d_n / \phi(n)] = 0$, and since $E[a_n d_n] \leq 1$, we have that

$$a_n d_n = (\phi(n+1) - \phi(n)) - \epsilon_n g(\phi(n))$$

is a martingale difference with respect to the filtration generated by market share. Note that $\sum_{n=1}^{\infty} a_n = \infty$ and $\sum_{n=1}^{\infty} a_n^2 < \infty$.

Furthermore, $C(\phi(n))$ is a probability vector, so $\|g(\phi(n))\| \leq 2$ and $\|\phi(n)(\cdot)\|_{\infty} \leq 1$, and $\|d_n\| \leq 2$.

Define $t_0 = 0$ and $t_n = \sum_{i=1}^n \frac{1}{i} \approx \log n$ for all n . Consider a continuous-time interpolation $\phi^0(\cdot)$ of $\phi(n)$ on $(-\infty, +\infty)$ by

$$\phi^0(t) = \begin{cases} \phi(0) & \text{for } t \leq 0, \\ \phi(n) & \text{for } t_n \leq t \leq t_{n+1}. \end{cases}$$

Also, define the sequence of the shifted process $\phi^n(\cdot)$ by

$$\phi^n(t) = \phi^0(t_n + t) \quad \text{for all } t \in (-\infty, +\infty). \quad (2.6)$$

The ODE method developed in stochastic approximation theory [Kushner and Yin 03] gives powerful tools for studying the behavior of this type of function sequence.

Next, we use [Kushner and Yin 03, Theorem 2.1]. This theorem implies that the sequence of functions $(\phi^n(\cdot))$ has a convergent subsequence such as $(\phi^{r_m}(\cdot))$, which converges uniformly to $\bar{\phi}(\cdot)$, the trajectory of the ODE

$$\frac{d}{dt}\bar{\phi}(t) = C(\bar{\phi}(t)) - \bar{\phi}(t).$$

Consider a sequence δ_m that converges to zero. Then there exists a subsequence $(\phi^{r_m}(\cdot))$ such that for all t , we have

$$\|\phi^{r_m}(t) - \bar{\phi}(t)\| \leq \delta_m.$$

From the definition, we know that $\phi^{r_m}(t) = \phi^0(t + t_{r_m}) = \phi(k)$ for some k such that $t_k \leq t_{r_m} + t < t_{k+1}$. So for each m and large enough k , we have

$$\|\phi(k) - \bar{\phi}(t_k - t_{r_m})\| \leq \delta_m.$$

Now we want to show that there exists a sequence $(\bar{t}_k)_{k \geq 1}$ such that as $\bar{t}_k \rightarrow \infty$,

$$\lim_{k \rightarrow \infty} |\phi(k) - \bar{\phi}(\bar{t}_k)| = 0,$$

where k is bounded by a constant that depends on m . Let us show this by choosing $k > u_m$, where u_m is a constant that depends on m . Let

$$f_m = \max(r_m^2 - 1, u_m).$$

Then f_m is an increasing sequence that goes to infinity. For each k such that $f_{m+1} > k \geq f_m$, define $\bar{t}_k = t_k - t_{r_m}$. It is clear that

$$\bar{t}_k = t_k - t_{r_m} > t_{r_m^2 - 1} - t_{r_m} > \sum_{i=r_m+1}^{r_m^2} \frac{1}{i} > \ln(r_m),$$

so \bar{t}_k goes to infinity as k increases. □

In Theorem 2.3, we have shown that the stochastic process of the market share asymptotically follows a deterministic path, which is the solution of an ODE. Before studying the stability of the equilibrium, let us give an informal interpretation of this theorem. Note that

$$\phi(n+1) - \phi(n) = \frac{C(\phi(n)) - \phi(n)}{n+1} + \frac{d_n}{n+1},$$

where $d_n = I^{n+1} - C(\phi(n))$ is a zero-mean random variable. As n goes to infinity, the last term on the right approaches zero. Further, the cumulative variance of the noise terms is bounded, since $\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$. Informally, this suggests that

the limiting behavior of the stochastic process $\phi(n)$ can be inferred from the limiting behavior of the deterministic process

$$\hat{\phi}(n+1) - \hat{\phi}(n) = \frac{C(\hat{\phi}(n)) - \hat{\phi}(n)}{n+1}.$$

With our next theorem we show that under some stronger assumptions, the market share sequence converges to an equilibrium.

Theorem 2.4. *Let $C : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be the choice probability function defined in (2.1), where h is the identity function, i.e.,*

$$C_i(X) = P \left(\operatorname{argmax}_j (J\phi_j + f_j + \epsilon_j) = i \right).$$

Also, assume that the random vector ϵ admits a strictly positive density on \mathbb{R}^m and is such that C is continuously differentiable. If $\bar{\phi}(t)$ is a trajectory of the ODE

$$\frac{d}{dt} \bar{\phi}(t) = C(\bar{\phi}(t)) - \bar{\phi}(t), \tag{2.7}$$

then $\bar{\phi}(t)$ is well defined on all of \mathbb{R} , and

$$\lim_{t \rightarrow \infty} (C(\bar{\phi}(t)) - \bar{\phi}(t)) = 0. \tag{2.8}$$

In order to prove the stability of the ODE, we use the stochastic Lyapunov function approach.

Proof. In [Hofbauer and Sandholm 02, Theorem 2.1], it is shown that the derivative matrix of $C(X)$ represented as $DC(X)$ is symmetric. Then the vector field C admits a convex potential function $\tilde{W} : \mathbb{R}^m \rightarrow \mathbb{R}$ such that $\nabla \tilde{W} = C$. Now let us define $W : \mathbb{R}^m \rightarrow \mathbb{R}$ as

$$W(P) = \frac{1}{2} (P'P) - \tilde{W}(P).$$

It is clear that for all $P \in \mathbb{R}^m$, we have

$$\nabla W(P) = (P - C(P))'.$$

So the domain of the trajectories of (2.8) is all of $(-\infty, +\infty)$. Let $\bar{\phi}(\cdot)$ be the trajectory of the ODE in (2.7). Then

$$\begin{aligned} \frac{d}{dt} W(\bar{\phi}(t)) &= (\nabla W(\bar{\phi}(t))) \left(\frac{d\bar{\phi}}{dt} \right) = (\bar{\phi}(t) - C(\bar{\phi}(t)))' (C(\bar{\phi}(t)) - \bar{\phi}(t)) \\ &= - \|\bar{\phi}(t) - C(\bar{\phi}(t))\|^2. \end{aligned}$$

The above equality holds if $\bar{\phi}(t) \in RP = \{p \in P^m \mid C(p) = p\}$.

In the next step, we will show that $|W(X)| \rightarrow \infty$ as $\|X\|_2 \rightarrow \infty$. Let $X = \alpha u$, where $\|u\|_2 = 1$. Clearly, $\|C(X)\|_2 \leq m$. We have

$$|\tilde{W}(X)| = \left| \int_0^\alpha d/dt(\tilde{W}(tu))dt \right| = \left| u \int_0^\alpha C(tu)dt \right| \leq \alpha m,$$

where the last step follows from the Cauchy–Schwarz inequality. So

$$W(X) = \frac{1}{2} \|X\|_2^2 - \tilde{W}(X)$$

goes to $+\infty$ as $\|X\|_2$ goes to infinity. Since $W(X)$ is smooth, the minimum of W over \mathbb{R}^m exists, and we have $\nabla W(X) = 0$ for all of its minimizers. So its minimum points belong to RP .

Furthermore, if $C(p) = p$, then p is an interior point of the simplex. If $p_i = 0$ for some i , then by definition,

$$C_i(p) = P \left(\operatorname{argmax}_j Jp + f_j + \epsilon_j = i \right) = 0.$$

This contradicts the assumption that for every j , the noise variable ϵ_j has positive density everywhere in \mathbb{R} . Hence, we can conclude that $W(X)$ is a Lyapunov function, and the trajectory of the ODE converges to a point belonging to the set RP . \square

First we showed that the market share converges to an equilibrium in our model. This equilibrium point belongs to the solution of $C(p) = p$, which does not necessarily have a unique solution. When $C(p) = p$ has multiple solutions, the market share can converge to one of these different points. In other words, we have several candidates for the equilibrium point, and the beginning behavior of the market will determine which one of these equilibria will be selected. So there is some $\bar{\phi} \in RP = \{p \in P^m \mid C(p) = p\}$ such that $\phi(n) \rightarrow \bar{\phi}$.

Before discussing the rate of convergence, let us give two examples of our model.

Example 2.5. (Simple example of size 3.) Suppose there are $m = 3$ products in the market of respective fitness $f_1 = 1$, $f_2 = 2$, and $f_3 = 3$.

When $J = 4$, there are three potential fixed points

$$CR = \{(0.9615, 0.0249, 0.0135), (0.0100, 0.8803, 0.1097), (0.0026, 0.0071, 0.9903)\}.$$

Running Monte Carlo simulation shows that the process converges to the first equilibrium with a probability close to 0.68 and converges to the second with

probability close to 0.3. When $J = 1$, there is only one fixed point:

$$CR = \{(0.0862, 0.2380, 0.675)\}.$$

Example 2.6. (Balls-in-bins process.) Our result also specializes to the *balls-in-bins process with feedback* studied in [Oliveira 08, Chung et al. 03] and discussed at the beginning of Section 2. In this model, balls are sequentially thrown into bins, with the probability that a bin with n balls gets the next ball is proportional to $f(n)$, for some feedback function f . The above references focus on the feedback function $f(n) = n^J$ for $J > 0$, and they show that this family of functions f has a phase transition at $J = 1$. Theorems 2.1 and 2.2 of [Chung et al. 03] state that for each bin i , $\bar{\phi}_i = \lim_{t \rightarrow \infty} \phi_i(t)$ exists. The authors show that when $J > 1$, monopoly occurs; that is, the fraction of balls in one of the bins goes to 1. If $J < 1$, then $\bar{\phi}_i = 1/m$ for all bins. And for $J = 1$, the limit vector $\bar{\phi}$ is uniformly distributed on the simplex. These results are corollaries of our Theorem 2.3. Let $h(x)$ be $\ln x$, let ϵ_j be independent and identically distributed random variables with the extreme value distribution, and suppose that all products have the same fitness. In this case, the probability choice function is

$$C_i(\phi) = \frac{\phi_i^J}{\sum_j \phi_j(t)^J},$$

and the ODE in (2.3) has the form

$$\frac{d}{dt} \phi_i(t) = \frac{\phi_i(t)^J}{\sum_j \phi_j(t)^J} - \phi_i(t).$$

It is easy to see that this dynamical system is stable, and the above results follow from Lemma 5.1.

Theorem 2.7. *Assume that $\phi(n)$ is a market-share sequence, where $C(\phi)$ is smooth, h is linear, and $\bar{\phi}_i$ is one of the isolated stable points of the ODE in (2.7). Then*

$$\lim_{n \rightarrow \infty} P \left\{ \sqrt{n}(\phi(n) - \bar{\phi}_i) < x \mid \lim_{n \rightarrow \infty} \phi(n) = \bar{\phi}_i \right\} = F_i(x),$$

where $F_i(x)$ stands for the Gaussian distribution with covariance matrix $\Sigma_1 = (DC(\bar{\phi}_i) - I/2)^{-1} \Sigma$, where $\Sigma = \text{diag}\{C(\bar{\phi}_i)\} - C^T(\bar{\phi}_i)C(\bar{\phi}_i)$.

In the proof of Theorem 2.7, first we use the standard analysis on the rate of convergence for general unconstrained stochastic approximations to show that when the equilibrium point is unique,

$$U^n = \sqrt{n}(\phi(n) - \bar{\phi})$$

converges weakly to a Gaussian distribution. So for large enough n , with high probability,

$$\|\phi(n) - \bar{\phi}\| \leq \frac{\epsilon}{\sqrt{n}}.$$

In other words, for a sufficiently large number of people in the market, with high probability, the difference between the market share and the equilibrium point is less than ϵ/\sqrt{n} , where ϵ is a function of the social influence parameter J . This theorem shows that the variance of the stationary measure is a function of the social influence. Next, by redefining the process, we show that this result also holds in the case of multiple equilibria.

Proof. Let $U^n = \sqrt{n}(\phi(n) - \bar{\phi})$. We will initially show that when there is a unique equilibrium point, U^n converges to a Gaussian distribution with covariance matrix Σ_1 .

Let $\bar{\phi}$ be the isolated stable point of the ODE in (2.3). As we have done in the proof of Theorem 2.3, while constructing $\phi^n(\cdot)$ from the sequence $\phi(n)$, let $U_n(\cdot)$ be a piecewise constant right-continuous interpolation of the sequence $\{U^i \mid i \geq n\}$ on $[0, \infty)$. So $U_n(t) = U^k$, where $t_k \leq t + t_n \leq t_{k+1}$.

Let $a_n = \frac{1}{n+1}$, so that $\phi(n)$ satisfies the equation

$$\phi(n+1) = \phi(n) + a_n g(\phi(n)) + a_n d_n = \phi(n) + a_n Y_n, \quad (2.9)$$

where $g(\phi(n)) = C(\phi) - \phi$ and $a_n d_n$ is a martingale difference. So $\phi(n)$ satisfies [Kushner and Yin 03, Algorithm 10.2.1]. One can show that the assumptions of [Kushner and Yin 03, Theorem 10.2.1] are satisfied (see Theorem 6.1), which gives us the following result: $U_n(\cdot)$ converges weakly in $D^m[0, \infty]$ to a stationary process $U(\cdot)$, where $U(\cdot)$ is the solution of

$$U(t) = U(0) + \int_0^t (DC(\bar{\phi}) - I/2)U(s)ds + W(t),$$

and where $W(\cdot)$ is a Wiener process with covariance matrix $\Sigma = \text{diag}\{C(\bar{\phi})\} - C^T(\bar{\phi})C(\bar{\phi})$.

We know that the solution of this stochastic differential equation is an Ornstein–Uhlenbeck process and that its stationary distribution is Gaussian with covariance matrix $\Sigma_1 = (DC(\bar{\phi}) - I/2)^{-1}\Sigma$.

In the case of multiple equilibrium points, given that the process converges to one of them, we can redefine the process such that it has a unique equilibrium:

$$\tilde{g}(x) = \begin{cases} g(x) & \text{if } |x - \bar{\phi}_i| < \epsilon_1, \\ h(x) & \text{if } \epsilon_1 < |x - \bar{\phi}_i| < \epsilon_2, \\ (x - \bar{\phi}_i) & \text{if } |x - \bar{\phi}_i| > \epsilon_2, \end{cases}$$

where $g(x) = C(x) - x$. The function $h(x)$ is defined such that $\tilde{g}(x)$ is a smooth function with a unique zero point.

Let

$$t = \inf\{n : |\phi_n - \bar{\phi}_i| < \epsilon_1, \forall s > n\}.$$

From Theorem 2.3, we know that $a_n d_n$ is a martingale difference with respect to $\tilde{\phi}_{(i)}$ and $\tilde{d}_n = I^{n+1} - C(\tilde{\phi}(n))$. So

$$\tilde{\phi}(n+1) = \tilde{\phi}(n) + \frac{1}{n+1} \tilde{g}(\tilde{\phi}(n)) + \frac{1}{n+1} \tilde{d}_n.$$

For $s \geq t$, we have $\phi_s = \tilde{\phi}_s$. Using the fact from Theorem 2.3 that the process converges, we obtain

$$\begin{aligned} & \lim_{m \rightarrow \infty} P\left\{\sqrt{m}(\phi(m) - \bar{\phi}_i) < x, \lim_{n \rightarrow \infty} \phi(n) = \bar{\phi}_i\right\} \\ &= \lim_{m \rightarrow \infty} \sum_{n=0}^{\infty} P\left\{\sqrt{m}(\phi(m) - \bar{\phi}_i) < x, t = n, \lim_{n \rightarrow \infty} \phi(n) = \bar{\phi}_i\right\} \\ &= \sum_{n=0}^{\infty} \lim_{m \rightarrow \infty} P\left\{\sqrt{m}(\phi(m) - \bar{\phi}_i) < x, t = n\right\} \\ &= \sum_{n=0}^{\infty} \lim_{m \rightarrow \infty} P\left\{\sqrt{m}(\tilde{\phi}(m) - \bar{\phi}_i) < x, t = n\right\} \\ &= \lim_{m \rightarrow \infty} \sum_{n=0}^{\infty} P\left\{\sqrt{m}(\tilde{\phi}(m) - \bar{\phi}_i) < x, t = n\right\} \\ &= \lim_{m \rightarrow \infty} \sum_{n=0}^{\infty} E\left\{I(\sqrt{m}(\tilde{\phi}(m) - \bar{\phi}_i) < x, t = n)\right\} \\ &= \lim_{m \rightarrow \infty} P\left\{\sqrt{m}(\tilde{\phi}(m) - \bar{\phi}_i) < x, t < \infty\right\}. \end{aligned}$$

The last step follows from the Lebesgue dominated convergence theorem, since the indicator function is less than 1. Let E_n be the event $E_n = \{|\phi_n - \bar{\phi}_i| < \delta\}$, where $\delta \in (0, 1)$. Since $\tilde{\phi}(m)$ has a unique equilibrium point, we can use our result from the first part:

$$\begin{aligned} & \lim_{m \rightarrow \infty} P\left\{\sqrt{m}(\tilde{\phi}(m) - \bar{\phi}_i) < x, E_n\right\} \\ &= \lim_{m \rightarrow \infty} E\left[P\left\{\sqrt{m}(\tilde{\phi}(m) - \bar{\phi}_i) < x \mid E_n\right\}\right] P(E_n) \\ &= F_i(x) P(E_n). \end{aligned}$$

Let Δ denote the symmetric difference. Then

$$\begin{aligned} & \left| P\left\{\sqrt{m}(\tilde{\phi}(m) - \bar{\phi}_i) < x, E_n\right\} - P\left\{\sqrt{m}(\tilde{\phi}(m) - \bar{\phi}_i) < x, \lim_{n \rightarrow \infty} \phi(n) = \bar{\phi}_i\right\} \right| \\ & \leq P\left\{E_n \Delta \left\{\lim_{n \rightarrow \infty} \phi(n) = \bar{\phi}_i\right\}\right\}. \end{aligned}$$

So as n goes to infinity, $P\left\{E_n \Delta \left\{\lim_{n \rightarrow \infty} \phi(n) = \bar{\phi}_i\right\}\right\} \rightarrow 0$. Thus, we conclude that

$$\lim_{n \rightarrow \infty} P\left\{\sqrt{n}(\phi(n) - \bar{\phi}_i) < x \mid \lim_{n \rightarrow \infty} \phi(n) = \bar{\phi}_i\right\} = F_i(x).$$

□

This result shows that the convergence rate is a function of the social influence parameter J . Also, if there are multiple equilibria, then the rates of convergence to each equilibrium point are different, since the covariance matrix Σ_1 depends on the equilibrium. One can think of the asymptotic behavior of the process as follows. One of the stable points is selected as a result of the earlier behavior of the process, and then a convergence mechanism starts and drives the process to this point.

3. A Special Case: The Logit Model

The most common choice probability function used to model herding is the *logit choice function* [Sorensen 06, Glaeser et al. 96]. Specifically, the multinomial logit model is the most commonly used parametric model in economics, operations management, and marketing [Anderson et al. 92]. If ϵ_j are independent identically distributed random variables with the extreme value distribution (Gumbel distribution) such that

$$F(\epsilon) = \exp(-\exp(-\eta^{-1}\epsilon - \gamma)),$$

where γ is Euler's constant, then the choice probability function is the logit choice function

$$L_i(X) = \frac{e^{\eta^{-1}x_i}}{\sum_j e^{\eta^{-1}x_j}}.$$

A major advantage of the multinomial logit model is that it is analytically tractable. In particular, it has a closed-form expression for the choice probabilities. In this section, we will focus on the case that the choice probability function is multinomial logit. As we have shown in the previous section, this is a special

case of our model

$$i \in \operatorname{argmax}_j (Jh(\phi_j(n)) + f_j + \epsilon_j)$$

in which the function h is the identity function and the noise terms ϵ_j are independent and identically Gumbel distributed.

By finding an approximate solution for the equilibrium of the multinomial logit model, we have the following three observations: First, in case of weak social influence, i.e., when J is small enough, there is a unique equilibrium. Second, with strong social influence, monopoly occurs. In this case, eventually any of the products can get the largest market share, so the number of equilibria is m . Third, inequality in the market increases with the social influence. By inequality, we mean the differences in market share among the various products. These results support the experimental result of [Salganik et al. 06].

Theorem 3.1. *For the logit choice function, there exists J^* such that the following hold:*

- *When $J \leq J^*$, there is a unique equilibrium*

$$\frac{e^{f_i}}{A + \sqrt{cJB}} + \frac{Je^{2f_i}}{(A + \sqrt{BJ})^2} \leq \bar{\phi}_i \leq \frac{e^{f_i}}{0.8A + \sqrt{\frac{1}{2}JB}} + \frac{Jce^{2f_i}}{(0.8A + \sqrt{\frac{1}{2}BJ})^2},$$

where A and B are constants.

- *When $J > J^*$, for each equilibrium, there exists some i such that $1 - \bar{\phi}_i = \Theta(e^{-J})$. In other words, product i has a monopoly in the market.*

We give here a sketch of the proof; for a detailed proof see Section 7. According to Theorem 2.4, equilibrium points satisfy

$$C_i(\bar{\phi}) = \frac{e^{J\bar{\phi}_i + f_i}}{\sum_{i=1}^m e^{J\bar{\phi}_i + f_i}} = \bar{\phi}_i, \quad \sum_{i=1}^m \bar{\phi}_i = 1, \quad -DC(\bar{\phi}) + I \succeq 0, \quad (3.1)$$

where $DC(\bar{\phi})$ denotes the derivative of $C(\bar{\phi})$ and \succeq stands for positive semidefinite.

Let $F(x) = x - \ln x$ and $u = \ln(\sum_{i=1}^m e^{J\bar{\phi}_i + f_i})$. From the first equation in (3.1), we have

$$e^{J\bar{\phi}_i + f_i} = e^u \bar{\phi}_i. \quad (3.2)$$

When we rearrange (3.2), we get

$$J\bar{\phi}_i - \ln(J\bar{\phi}_i) = u - f_i - \ln J.$$

Then

$$F(J\bar{\phi}_i) = u - f_i - \ln J. \quad (3.3)$$

We can define two inverse functions for F , $F_1^{-1}(x)$ and $F_2^{-1}(x)$, such that $F_1^{-1}(x) > 1$ and $F_2^{-1}(x) \leq 1$. Then $F_1^{-1}(x)$ is an increasing function, while $F_2^{-1}(x)$ is a decreasing function. Thus, (3.3) can be written as

$$\bar{\phi}_i = \frac{1}{J} F_{1 \text{ or } 2}^{-1}(u - f_i - \ln J). \quad (3.4)$$

Let $S = \{i : \bar{\phi}_i \leq 1/J\}$. Therefore, using (3.4), we see that (3.1) is equivalent to finding u such that

$$\sum_{i \in S} \frac{1}{J} F_2^{-1}(u - f_i - \ln J) + \sum_{i \notin S} \frac{1}{J} F_1^{-1}(u - f_i - \ln J) = 1. \quad (3.5)$$

From the fact that $-DC(\bar{\phi}) + I$ is positive semidefinite, we can conclude that $|S| \geq m - 1$, i.e., that there exists at most one i such that $\bar{\phi}_i \geq 1/J$.

Let $\bar{f} = \max_i f_i$ and

$$J^* = \sum_i F_2^{-1}(\bar{f} - f_i + 1).$$

Equation (3.5) has a unique solution when $J \leq J^*$ if and only if $S = \{1, \dots, m\}$. We refer to this case as the weak social influence case.

On the other hand, when $J > J^*$, (3.5) has a solution when there exists an i such that $i \notin S$. We show that in this case, in equilibrium, $\bar{\phi}_i$ is close to 1, and all other $\bar{\phi}_j$ are close to zero. We call this the strong social influence case.

In both of these cases, we can use the following inequalities to find an approximation of the equilibrium points:

$$x + \ln x + \frac{1}{2} \geq F_1^{-1}(x) \geq x + \ln x, \quad e^{-x} e^{e^{-x+1}} \geq F_2^{-1}(x) \geq e^{-x} e^{e^{-x}}. \quad (3.6)$$

By applying (3.6), we have the following bounds for the weak social influence case:

$$\bar{\phi}_i \leq \frac{e^{f_i}}{0.8A + \sqrt{\frac{1}{2}JB}} + \frac{Jce^{2f_i}}{\left(0.8A + \sqrt{\frac{1}{2}BJ}\right)^2}, \quad \bar{\phi}_i \geq \frac{e^{f_i}}{A + \sqrt{cJB}} + \frac{Je^{2f_i}}{(A + \sqrt{BJ})^2}.$$

In the strong social influence case, using (3.6), we can find a bound for $\bar{\phi}_i$ when $i \notin S$:

$$A_i e^{-J+\alpha} + cB_i J e^{-2(J-\alpha)} \geq 1 - \bar{\phi}_i \geq A_i e^{-J} + B_i J e^{-2J}.$$

4. Conclusions

We have presented an analysis for the dynamics of market share when social influence is present. Using techniques from stochastic approximation theory that relate the limit behavior of a stochastic process to the limit behavior of a differential equation, we show that market share converges to an equilibrium in our model. Our result also implies that there may exist several different candidates for the eventual market share. As the degree of social influence increases, so does the size of the set of equilibria for the market share.

We also analyze the rate of convergence of the equilibrium. We prove that the interpolated process formed by the centered and normalized iterate converges weakly to a Gaussian diffusion. In other words, for a sufficiently large number n of people in the market, with high probability, the difference between the market share and the equilibrium point is less than c/\sqrt{n} , where c depends on the social influence. Our result also shows that in case of multiple equilibria, the convergence rate to each of them is different.

We also focus on the case that the choice model is the multinomial logit model. In this special case, we show that inequality in the market increases with social influence, and with large enough social influence, monopoly occurs. Our observations for the logit model support the empirical results of a study from a recent web-based music market experiment [Salganik et al. 06].

5. Appendix A: Lemma 5.1

Lemma 5.1. *Let*

$$\frac{X_i(t)^J}{\sum_j X_j(t)^J} = \frac{d}{dt} X_i(t) + X_i(t).$$

Then this ODE is stable, and the equilibrium points are as follows: If $J > 1$, monopoly occurs, i.e., $(1, 0, 0, \dots, 0)$ is the fixed point. If $J < 1$, all products have the same market share, i.e., $(1/m, \dots, 1/m)$ is the fixed point.

Proof. The equality

$$\frac{X_i(t)^J}{\sum_i X_i(t)^J} = \frac{d}{dt} X_i(t) + X_i(t)$$

implies that

$$\frac{1}{\sum_j X_j(t)^J} = X_i(t)^{-J} \left(\frac{d}{dt} X_i(t) + X_i(t) \right)$$

for all i . Then

$$X_i(t)^{-J} \frac{d}{dt} X_i(t) + X_i(t)^{-J+1} = X_j(t)^{-J} \frac{d}{dt} X_j(t) + X_j(t)^{-J+1}.$$

Therefore,

$$\begin{aligned} \frac{d}{dt} e^{(-J+1)t} X_i(t)^{-J+1} &= \frac{d}{dt} e^{(-J+1)t} X_j(t)^{-J+1}, \\ e^{(-J+1)t} X_i(t)^{-J+1} - X_i(0)^{-J+1} &= e^{(-J+1)t} X_j(t)^{-J+1} - X_j(0)^{-J+1}, \end{aligned}$$

and

$$X_i(t)^{-J+1} - X_j(t)^{-J+1} = e^{(J-1)t} (X_i(0)^{-J+1} - X_j(0)^{-J+1}). \quad (5.1)$$

When $X_i(0) \neq X_j(0)$, from (5.1), if $J > 1$, the left-hand side goes to infinity, and if $J < 1$, the left-hand side goes to zero. \square

6. Appendix B: Theorem 6.1

Theorem 6.1. *Assume that $\phi(n)$ is a market-share sequence, where $C(\phi)$ is smooth and $\bar{\phi}$ is an isolated stable point of the ODE in (2.3). Then there is some matrix Σ_1 such that*

$$U^n = \sqrt{n} (\phi(n) - \bar{\phi})$$

converges weakly to the normal distribution with covariance Σ_1 .

Proof. Let $U^n = \sqrt{n}(\phi(n) - \bar{\phi})$. As we have done in the proof of Theorem 2.3, while constructing $\phi^n(\cdot)$ from the sequence $\phi(n)$, let $U_n(\cdot)$ be a piecewise-constant right-continuous interpolation of the sequence $\{U^i \mid i \geq n\}$ on $[0, \infty)$. So $U_n(t) = U^k$, where $t_k \leq t + t_n \leq t_{k+1}$.

Let $a_n = \frac{1}{n+1}$, so that $\phi(n)$ satisfies the equation

$$\phi(n+1) = \phi(n) + a_n g(\phi(n)) + a_n d_n = \phi(n) + a_n Y_n, \quad (6.1)$$

where $g(\phi(n)) = C(\phi) - \phi$, and $a_n d_n$ is a Martingale difference. So $\phi(n)$ satisfies [Kushner and Yin 03, Algorithm 10.2.1]. We will use [Kushner and Yin 03, Theorem 10.2.1] for our proof, so let us check the assumptions of this theorem, A2.0 to A2.7, as presented in [Kushner and Yin 03].

Assumption A2.0 holds because $a_n = \frac{1}{n+1}$.

We have already shown in the proof of Theorem 2.3 that

$$g(\phi(n)) + d_n = I^{n+1} - \phi(n) \quad \text{and} \quad \|g(\phi(n))\|_2 < 1,$$

so $\|d_n\|_2 < 2$. Thus,

$$\|Y_n\| = \|g(\phi(n)) + d_n\| \leq 2.$$

As a result, $\{Y_n I_{|\phi(n) - \bar{\phi}| \leq \rho}\}$ is uniformly integrable for small $\rho > 0$, which satisfies A.2.1. As already shown, $\phi(n) \rightarrow \bar{\phi}$, so A.2.2 holds. The tightness condition A.2.3 can be concluded from [Kushner and Yin 03, Theorem 10.4.1].

By the assumption of C being smooth, $g(\phi(n)) = C(\phi) - \phi$ is also smooth, so assumptions A.2.4 and A.2.6 hold, where in our case $A = DC(\phi) - I$ is a Hurwitz matrix.

Since $g(\bar{\phi}(n)) = C(\bar{\phi}) - \bar{\phi} = 0$, assumption A.2.5 is satisfied.

Define $\Sigma = \text{diag}\{C(\bar{\phi})\} - C^T(\bar{\phi})C(\bar{\phi})$. From $d_n = I^{n+1} - C(\phi(n))$ and $\|d_n\|_2 < 2$, it can be shown that for small $\rho > 0$,

$$E_n[(d_n)(d_n)' I_{\{|\phi(n) - \bar{\phi}| \leq \rho\}}] \rightarrow \Sigma.$$

Thus A.2.7 holds.

Now we can apply [Kushner and Yin 03, Theorem 10.2.1] to see that $U_n(\cdot)$ converges weakly in $D^m[0, \infty]$ to a stationary process $U(\cdot)$, where $U(\cdot)$ is the solution of

$$U(t) = U(0) + \int_0^t (A + I/2)U(s)ds + W(t),$$

where $W(\cdot)$ is a Wiener process with covariance matrix Σ . We know that the solution of this stochastic differential equation is an Ornstein–Uhlenbeck process and that its stationary distribution is Gaussian with covariance matrix $\Sigma_1 = (A + I/2)^{-1}\Sigma$. □

7. Appendix C: Detailed Proof of Theorem 3.1

Theorem 7.1. *For the logit choice function, there exists J^* such that the following hold:*

- *When $J \leq J^*$, there is a unique equilibrium*

$$\frac{e^{f_i}}{A + \sqrt{cJB}} + \frac{J e^{2f_i}}{(A + \sqrt{BJ})^2} \leq \bar{\phi}_i \leq \frac{e^{f_i}}{0.8A + \sqrt{\frac{1}{2}JB}} + \frac{J c e^{2f_i}}{\left(0.8A + \sqrt{\frac{1}{2}BJ}\right)^2},$$

where A and B are constants.

- When $J > J^*$, for each equilibrium, there exists some i such that $1 - \bar{\phi}_i = \Theta(e^{-J})$. In other words, product i has a monopoly in the market.

Proof. Finding the solution of $C(\bar{\phi}) = \bar{\phi}$, where $-DC(\bar{\phi}) + I$ is positive semidefinite, for the logit case is equivalent to solving the following:

$$C_i(\bar{\phi}) = \frac{e^{J\bar{\phi}_i + f_i}}{\sum_{i=1}^m e^{J\bar{\phi}_i + f_i}} = \bar{\phi}_i, \quad \sum_{i=1}^m \bar{\phi}_i = 1, \quad -DC(\bar{\phi}) + I \succeq 0. \quad (7.1)$$

Let $u = \ln(\sum_{i=1}^m e^{J\bar{\phi}_i + f_i})$. Then from the first equation in (7.1), we obtain

$$e^{J\bar{\phi}_i + f_i} = e^u \bar{\phi}_i. \quad (7.2)$$

When we rearrange (7.2), we get

$$J\bar{\phi}_i - \ln(J\bar{\phi}_i) = u - f_i - \ln J.$$

Then

$$F(J\bar{\phi}_i) = u - f_i - \ln J, \quad (7.3)$$

where $F(x) = x - \ln(x)$. We can define two inverse functions for F , $F_1^{-1}(x)$ and $F_2^{-1}(x)$, such that $F_1^{-1}(x) > 1$ and $F_2^{-1}(x) \leq 1$.

Since

$$(F_1^{-1})'(x) = \frac{1}{1 - F_1^{-1}(x)} \geq 0,$$

it follows that $F_1^{-1}(x)$ is an increasing function, and similarly, $F_2^{-1}(x)$ is a decreasing function. From (7.3) we can conclude that

$$\bar{\phi}_i = \frac{1}{J} F_{1 \text{ or } 2}^{-1}(u - f_i - \ln J). \quad (7.4)$$

We can bound the above inverse functions as shown in Lemma 7.2:

$$x + \ln x + \frac{1}{2} \geq F_1^{-1}(x) \geq x + \ln x,$$

$$e^{-x} e^{e^{-x+1}} \geq F_2^{-1}(x) \geq e^{-x} e^{e^{-x}}.$$

Suppose $\bar{\phi} = (\bar{\phi}_1, \dots, \bar{\phi}_n)$ is the solution of (7.1) and let $S = \{i \mid \bar{\phi}_i \leq 1/J\}$. Combining (7.1) and (7.4), we have

$$\sum_{i \in S} \frac{1}{J} F_2^{-1}(u - f_i - \ln J) + \sum_{i \notin S} \frac{1}{J} F_1^{-1}(u - f_i - \ln J) = 1. \quad (7.5)$$

Case 1: Let us look first at the case that $S = \{1, \dots, n\}$, in other words, that $\bar{\phi}_i \leq 1/J$ for all i . From Lemma 7.4, this is equivalent to the weak social influence case, in which $J \leq J^*$. Then from (7.5), we have

$$\sum_i \frac{1}{J} F_2^{-1}(u - f_i - \ln J) = 1.$$

In this case, there is a unique equilibrium, since F_2^{-1} is decreasing, and the above equation has a solution if and only if $J \leq J^*$; see Lemma 7.4. Next, we will find an approximate solution for this equilibrium. From Lemma 7.2, we have the following bound for F_2^{-1} :

$$e^{-x} e^{e^{-x+1}} \geq F_2^{-1}(x) \geq e^{-x} e^{e^{-x}}.$$

So

$$\sum_i \frac{1}{J} e^{-(u-f_i-\ln J)} e^{e^{-(u-f_i-\ln J)+1}} \geq 1, \quad \sum_i \frac{1}{J} e^{-(u-f_i-\ln J)} e^{e^{-(u-f_i-\ln J)}} \leq 1.$$

From Lemma 7.3,

$$e^{-u} \sum_i e^{f_i} (1 + J e^{-u} e^{f_i}) \leq 1, \quad e^{-u} \sum_i e^{f_i} (1 + cJ e^{-u} e^{f_i}) \geq 1.$$

Thus

$$\left(\sum_i e^{f_i} \right) e^{-u} + e^{-2u} J \sum_i e^{2f_i} \leq 1, \quad \left(\sum_i e^{f_i} \right) e^{-u} + e^{-2u} cJ \sum_i e^{2f_i} \geq 1.$$

Let $A = \sum_i e^{f_i}$ and $B = \sum_i e^{2f_i}$. Then

$$Ae^{-u} + JB e^{-2u} \leq 1, \quad Ae^{-u} + JBc e^{-2u} \geq 1.$$

From rearranging the above, we get two quadratic equations. Using their roots, we can find lower and upper bounds on e^u , which we call e^{u_l} and e^{u_u} , respectively:

$$e^{2u} - Ae^u - JB \geq 0 \Rightarrow e^u \geq \frac{A}{2} + \sqrt{\frac{A^2}{4} + JB} = e^{u_l},$$

$$e^{2u} - Ae^u - JBc \leq 0 \Rightarrow e^u \leq \frac{A}{2} + \sqrt{\frac{A^2}{4} + JBc} = e^{u_u}.$$

From these inequalities and Lemma 7.2, we obtain

$$u_l \leq u \leq u_u \quad \text{and} \quad e^{-u_u+f_i} e^{e^{u_u+f_i} J} \leq \bar{\phi}_i \leq e^{-u_l+f_i} e^{e^{-u_l+f_i+1} J}.$$

Also, we can bound e^{u_l} and e^{u_u} as follows:

$$e^{u_l} \geq 0.8A + \sqrt{\frac{1}{2}JB}, \quad e^{u_u} \leq A + \sqrt{JBc}. \quad (7.6)$$

By applying Lemma 7.3 and using (7.6), we can have a simpler but weaker bound:

$$\bar{\phi}_i \leq \frac{e^{f_i}}{0.8A + \sqrt{\frac{1}{2}JB}} + \frac{Jce^{2f_i}}{\left(0.8A + \sqrt{\frac{1}{2}BJ}\right)^2}, \quad \bar{\phi}_i \geq \frac{e^{f_i}}{A + \sqrt{cJB}} + \frac{Je^{2f_i}}{(A + \sqrt{BJ})^2},$$

where $c = e/(1 - e^{-1})$.

Case 2: Let us look at the case $S \neq \{1, \dots, n\}$, in other words, the case in which there exists at least one i such that $\bar{\phi}_i > 1/J$. Next, we will show that there is at most one $\bar{\phi}_i$ for which $\bar{\phi}_i > 1/J$.

In the proof of Theorem 2.4, we showed that the equilibrium point $\bar{\phi}$ is the minimum of $W(\bar{\phi})$ over \mathbb{R}^m , where $\nabla W(\bar{\phi}(t)) = \bar{\phi} - C(\bar{\phi})$. So $\nabla^2 W(\bar{\phi}(t)) = -DC(\bar{\phi}) + I$ is positive semidefinite at the point $\bar{\phi}$:

$$A = -DC(\bar{\phi}) + I \succeq 0.$$

If we look at the diagonal elements of A , then

$$\begin{aligned} a_{ii} &= -\frac{Je^{J\bar{\phi}_i+f_i} \left(\sum_i e^{J\bar{\phi}_i+f_i}\right) - e^{J\bar{\phi}_i+f_i} \left(Je^{J\bar{\phi}_i+f_i}\right)}{\left(\sum_i e^{J\bar{\phi}_i+f_i}\right)^2} + 1 \\ &= -J \frac{e^{J\bar{\phi}_i+f_i}}{\sum_i e^{J\bar{\phi}_i+f_i}} + J \left(\frac{e^{J\bar{\phi}_i+f_i}}{\sum_i e^{J\bar{\phi}_i+f_i}}\right)^2 + 1 \\ &= -J\bar{\phi}_i + J\bar{\phi}_i^2 + 1. \end{aligned}$$

Thus $a_{ii} = 1 - J\bar{\phi}_i(1 - \bar{\phi}_i)$, and since $a_{ii} \geq 0$, we have that $J\bar{\phi}_i(1 - \bar{\phi}_i) \leq 1$.

So $0 \leq \bar{\phi}_i^2 - \bar{\phi}_i + 1/J$, which gives us for $J \geq 2$,

$$\bar{\phi}_i \geq \frac{1}{2} + \sqrt{\frac{1}{2} - \frac{1}{J}},$$

or

$$\bar{\phi}_i \leq \frac{1}{2} - \sqrt{\frac{1}{2} - \frac{1}{J}} \leq \frac{1}{J}.$$

From the above inequalities we can conclude that if $\bar{\phi}_i > 1/J$, then $\bar{\phi}_i > 1/2$.

Since $\sum_{i=1}^m \bar{\phi}_i = 1$, the above inequalities show that there can be at most one $\bar{\phi}_i$ for which $\bar{\phi}_i > 1/J$ if $J \geq 2$. The same argument follows from $\sum_{i=1}^m \bar{\phi}_i = 1$ when $J < 2$.

Now we consider the case that for exactly one product, $\bar{\phi}_i > 1/J$, and for all the other $j \neq i$, $\bar{\phi}_j \leq 1/J$. Therefore, $S = \{1, \dots, n\}/i$ and

$$\frac{1}{J}F_1^{-1}(u - f_i - \ln J) + \sum_{j \neq i} \frac{1}{J}F_2^{-1}(u - f_j - \ln J) = 1. \quad (7.7)$$

From Lemma 7.2 and the fact that $\bar{\phi}_i > \frac{1}{2}$, we obtain

$$\begin{aligned} \bar{\phi}_i &= \frac{1}{J} F_1^{-1}(u - f_i - \ln J) > \frac{1}{2}, \\ u - f_i - \ln J &> F\left(\frac{J}{2}\right) = \frac{J}{2} - \log\left(\frac{J}{2}\right) u > f_i + \frac{J}{2} + \log 2. \end{aligned}$$

Also,

$$1 \geq \bar{\phi}_i = \frac{1}{J} F_1^{-1}(u - f_i - \ln J),$$

and therefore $u \leq f_i + J$. Now define

$$D^i(u) = \frac{1}{J} \sum_{j \neq i} \frac{1}{J} F_2^{-1}(u - f_j - \ln J). \tag{7.8}$$

According to Lemmas 7.2 and 7.3, we have

$$e^{-x}(1 + ce^{-x}) \geq e^{-x} e^{e^{-x+1}} \geq F_2^{-1} \geq e^{-x} e^{e^{-x}} \geq e^{-x}(1 + e^{-x}).$$

Therefore,

$$\begin{aligned} D^i(u) &\geq \sum_{j \neq i} e^{-u+f_j} (1 + J e^{-u+f_j}) \geq \sum_{j \neq i} e^{-J+f_j-f_i} (1 + J e^{-J+f_j-f_i}) \\ &\geq e^{-J} A_i + \frac{e^{-2J}}{J} B_i, \end{aligned}$$

where $A_i = \sum_{j \neq i} e^{f_j-f_i}$ and $B_i = \sum_{j \neq i} e^{2f_j-2f_i}$. Similarly, we can find an upper bound for $D^i(u)$:

$$\begin{aligned} D^i(u) &\leq \sum_{j \neq i} e^{-u+f_j+\log J} (1 + cJ e^{-u+f_j+\log J}) \\ &\leq \sum_{j \neq i} \frac{1}{2} e^{-J/2+f_j-f_i} \left(1 + \frac{c}{2} J e^{-J/2+f_j-f_i}\right) \\ &\leq \frac{1}{2} e^{-J/2} A_i + \frac{c}{2} B_i \frac{e^{-J}}{J}. \end{aligned}$$

By replacing this inequality in (7.7), we get a tight approximation of u :

$$\frac{1}{J} F_1^{-1}(u - f_i - \log J) + \frac{1}{2} e^{-J/2} A_i + \frac{c}{2} B_i \frac{e^{-J}}{J} \geq 1.$$

When we apply F to both sides, we obtain

$$u - f_i - \log(J) \geq F\left(J - \frac{1}{2} A_i e^{-J/2} - \frac{c}{2} B_i \frac{e^{-J}}{J}\right).$$

Since $F(x) = x - \ln(x)$, we have

$$u \geq f_i + J - \left(\frac{1}{2}A_i e^{-J/2} + \frac{c}{2}B_i \frac{e^{-J}}{J} \right) - \log \left(1 - \frac{1}{2}A_i e^{-J/2} - \frac{c}{2}B_i \frac{e^{-J}}{J} \right). \quad (7.9)$$

There exists some α such that for all $J > J^*$, the following inequality holds:

$$\alpha \geq \left(\frac{1}{2}A_i e^{-J/2} + \frac{c}{2}B_i \frac{e^{-J}}{J} \right) + \log \left(1 - \frac{1}{2}A_i e^{-J/2} - \frac{c}{2}B_i \frac{e^{-J}}{J} \right).$$

Then from (7.9), we can conclude that

$$f_i + J \geq u \geq f_i + J - \alpha.$$

By applying Lemma 7.2 to (7.4) and using the bounds above, for $j \neq i$, we have

$$e^{-f_i+f_j-J+\alpha} e^{J e^{-f_i+f_j-J+\alpha+1}} \geq \phi_j \geq e^{-f_i+f_j-J} e^{J e^{-f_i+f_j-J+\alpha}}.$$

From Lemma 7.3 we can find a weaker but simpler bound for ϕ_j :

$$\begin{aligned} e^{-f_i+f_j-J+\alpha} + cJ e^{2(-f_i+f_j-J+\alpha)} &\geq \phi_j \geq e^{-f_i+f_j-J} + cJ e^{2(-f_i+f_j-J)}, \\ A_i e^{-J+\alpha} + cB_i J e^{-2(J-\alpha)} &\geq 1 - \phi_i \geq A_i e^{-J} + B_i J e^{-2J}. \end{aligned}$$

□

Lemma 7.2. *Let $F(x) = x - \ln(x)$. One can find two inverse functions for F such that $F_1^{-1}(x)$ for $x \geq 1$ and $F_2^{-1}(x)$ for $x \leq 1$. Then*

$$\begin{aligned} x + \ln x + \frac{1}{2} &\geq F_1^{-1}(x) \geq x + \ln x, \\ e^{-x} e^{e^{-x+1}} &\geq F_2^{-1}(x) \geq e^{-x} e^{e^{-x}}. \end{aligned}$$

Proof. Let us establish the above inequalities in four parts.

(1) $x + \ln x + \frac{1}{2} \geq F_1^{-1}(x)$: if $\alpha \geq \frac{1}{2}$, then $e^\alpha \geq (1 - e^{-1})^{-1}$, $g(x) = (e^\alpha - 1)x - \ln x$ is convex, and $g(x) \geq 0$. Then we have

$$\begin{aligned} (e^\alpha - 1)x &\geq \alpha + \ln x, & e^\alpha x &\geq x + \alpha + \ln x, \\ \alpha + \ln x &\geq \ln(x + \alpha + \ln x), & x + \alpha + \ln x - \ln(x + \alpha + \ln x) &\geq x, \\ F(x + \alpha + \ln x) &\geq x, & x + \alpha + \ln x &\geq F_1^{-1}(x). \end{aligned}$$

(2) $F_1^{-1}(x) \geq x + \ln x$: In this case, $\ln x \leq \ln(x + \ln x)$ for $x \geq 1$:

$$F(x + \ln x) = x + \ln x - \ln(x + \ln x) \leq x \quad \text{and} \quad x + \ln x \leq F_1^{-1}(x).$$

(3) $F_2^{-1}(x) \geq e^{-x} e^{e^{-x}}$: Here we have $x \leq x + (e^{e^{-x}} - 1)e^{-x}$ for $x \geq 1$:

$$\begin{aligned} x &\leq x - e^{-x} + e^{e^{-x}} e^{-x} = e^{e^{-x}} e^{-x} - \ln(e^{e^{-x}} e^{-x}) = F(e^{e^{-x}} e^{-x}), \\ F_2^{-1}(x) &\geq e^{-x} e^{e^{-x}}. \end{aligned}$$

(4) $e^{-x} e^{e^{-x+1}} \geq F_2^{-1}(x)$: For all $x \geq 1$, $e^{-x} \leq e^{-1} \leq \frac{\ln e}{e}$. Then $e^{e^{-x}} \leq e \Rightarrow$

$$\begin{aligned} e^{-x} e^{e^{-x+1}} &\leq e^{-x+1}, \\ e^{-x} e^{e^{-x+1}} + x - e^{-x+1} &\leq x e^{-x} e^{e^{-x+1}} - (-x + e^{-x+1}) \leq x, \\ F(e^{-x} e^{e^{-x+1}}) &\leq x, \\ e^{-x} e^{e^{-x+1}} &\geq F_2^{-1}(x). \end{aligned}$$

□

Lemma 7.3. For $x \geq 1$,

$$1 + ce^{-(x+1)} \geq e^{e^{-x}} \geq 1 + e^{-x},$$

where $c = \frac{e}{1-e^{-1}}$.

Proof. Let $x' = e^{-x}$. Then

$$e^{e^{-x}} = e^{x'} = 1 + x' + \frac{x'^2}{2} + \dots \leq 1 + x'(1 + x' + \dots) \leq 1 + \frac{x'}{1-x'} \leq 1 + cx'.$$

For $x \geq 1$, we have $x' \leq e^{-1}$, and therefore $c = \frac{e}{1-e^{-1}}$. □

Lemma 7.4. Let $\bar{f} = \max_i f_i$ and

$$J^* = \sum_i F_2^{-1}(\bar{f} - f_i + 1),$$

and let $G(u) = \sum_{i \in S} \frac{1}{J} F_2^{-1}(u - f_i - \ln J)$. Then $G(u) = 1$ has a solution if and only if $J \leq J^*$.

Proof. Let us first show that if $S = \{1, \dots, m\}$, in other words, $\bar{\phi}_i \leq 1/J$ for all i , then $J \leq J^*$. For all $1 \leq i \leq m$, we have

$$u - f_i - \ln J = F(J\bar{\phi}_i) \geq 1,$$

since $F(x) = x - \ln x$. Then for all i ,

$$\begin{aligned} u - \ln J - 1 &\geq f_i, \quad u - \ln J - 1 \geq \bar{f}, \\ u - f_i - \ln J &\geq \bar{f} + \ln J + 1 - f_i - \ln J \geq \bar{f} - f_i + 1. \end{aligned}$$

Thus

$$J^* = \sum_{i \in S} \frac{1}{J} F_2^{-1}(\bar{f} - f_i + 1) \geq \sum_{i \in S} \frac{1}{J} F_2^{-1}(u - f_i - \ln J) = J.$$

From the definition of G and J^* , we have $G(f_i + \ln J + 1) = J^*/J \geq 1$. Since $\lim_{u \rightarrow \infty} G(u) = 0$ for some u , it follows that $G(u) = 1$. \square

Acknowledgments. Simla Ceyhan's research is supported by a William R. and Sara Hart Kimball Stanford Graduate Fellowship.

References

- [Acemoglu et al. 08] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar. "Bayesian Learning in Social Networks." NBER Working Paper No. W14040. Available online (<http://ssrn.com/abstract=1139356>), To appear in *Review of Economic Studies*, 2008.
- [Acemoglu et al. 10] D. Acemoglu, K. Bimpikis, and A. Ozdaglar. "Dynamics of Information Exchange in Endogenous Social Networks." To appear, 2010.
- [Anderson et al. 92] P. Anderson, J. de Palma, and F. Thisse. *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press, 1992.
- [Baddeley 07] M. Baddeley. "Analysing Herding: Insights from Neuroeconomics and Social Psychology." *Queens' Economics Seminar, Cambridge University*, 2007.
- [Banerjee 92] A. Banerjee. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics* 107:3 (1992), 797–817.
- [Ben-Akiva and Lerman 85] M. E. Ben-Akiva and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press, 1985.
- [Bertrand et al. 00] M. Bertrand, E. Luttmer, and S. Mullainathan. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics* 115:3 (2000), 1019–1055.
- [Bikhchandani et al. 92] S. Bikhchandani, D. Hirschleifer, and I. Welch. "A Theory of Fads, Fashions, Custom and Cultural Change as Informational Cascades." *Journal of Political Economy* 100:5 (1992), 992–1026.
- [Borgs et al. 07] C. Borgs, J. Chayes, C. Daskalakis, and S. Roch. "First to Market Is Not Everything: An Analysis of Preferential Attachment with Fitness." *Proceedings of ACM STOC 2007* pp. 135–144, 2007.
- [Cha et al. 07] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon. "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System." In *IMC '07: Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, pp. 1–14, 2007.
- [Chamley 03] C. Chamley. *Rational Herds—Economic Models of Social Learning*. Cambridge, UK: Cambridge University Press, 2003.
- [Chandukala et al. 08] S. Chandukala, J. Kim, and T. Otter. *Choice Models in Marketing*. Delft: Now Publishers Inc., 2008.

- [Chung et al. 03] F. Chung, S. Handjani, and D. Jungreis. “Generalizations of Pólya’s Urn Problem.” *Annals of Combinatorics* 7 (2003), 141–153.
- [Cialdini 01] R. Cialdini. *Influence: Science and Practice*. Needham Heights: Allyn and Bacon, 2001.
- [Dubin 95] R. Dubin. “Estimating Logit Models with Spatial Dependence.” In *New Directions in Spatial Econometrics*, pp. 229–242. Berlin: Springer-Verlag, 1995.
- [Ellison and Fudenberg 93] G. Ellison and D. Fudenberg. “Rules of Thumb for Social Learning.” *Journal of Political Economy* 101:4 (1993), 612–643.
- [Follmer 74] H. Follmer. “Random Economies with Many Interacting Agents.” *J. Math. Econ.* 1:51 (1974), 41–62.
- [Glaeser et al. 96] E. Glaeser, B. Sacerdote, and J. Scheinkman. “Crime and Social Interactions.” *Quarterly Journal of Economics* 111:2 (1996), 507–548.
- [Grossman and Stiglitz 76] S. Grossman and J. Stiglitz. “Information and Competitive Price Systems.” *American Economic Review* 66 (1976), 246–253.
- [Hofbauer and Sandholm 02] J. Hofbauer and W. Sandholm. “On the Global Convergence of Stochastic Fictitious Play.” *Econometrica* 70:6 (2002), 2265–2294.
- [Hong et al. 04] H. Hong, J. Kubik, and J. Stein. “Social Interaction and Stock Market Participation.” *Journal of Finance* 59 (2004), 137–163.
- [Kushner and Yin 03] J. Kushner and G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*, Stochastic Modelling and Applied Probability. New York: Springer, 2003.
- [Manski 00] C. Manski. “Economic Analysis of Social Interactions.” *Journal of Economic Perspectives* 14:3 (2000), 115–136.
- [McFadden 80] D. McFadden. “Econometric Models for Probabilistic Choice among Products.” *Journal of Business* 53:3 (1980), 13–29.
- [Montanari and Saberi 08] A. Montanari and A. Saberi. “Convergence to Equilibrium in Local Interaction Games.” In *Proceedings of the 50th Annual Symposium on Foundations of Computer Science (FOCS ’09)*, 2009.
- [Ost et al. 08] J. Oh, A. Susarla, and Y. Tan. “Examining the Diffusion of User-Generated Content in Online Social Networks.” Available online (<http://ssrn.com/abstract=1182631>), 2008.
- [Oliveira 08] R. Oliveira. “Balls-in-Bins Processes with Feedback and Brownian Motion.” *Combinatorics, Probability and Computing* 17 (2008), 87–110.
- [Paez et al. 08] A. Paez, D. Scott, and E. Volz. “A Discrete-Choice Approach to Modeling Social Influence on Individual Decision Making.” *Environment and Planning B: Planning and Design* 35:6 (2008), 1055–1069.
- [Ramaprasad and Dewan 09] J. Ramaprasad and S. Dewan. “Consumer Choice in an Online Music Community: Bandwagon Effects and Local Network Influence.” Manuscript, 2009.
- [Ratliff et al. 08] R. M. Ratliff, V. Rao, C. P. Narayan, and K. Yellepeddi. “A Multi-flight Recapture Heuristic for Estimating Unconstrained Demand from Airline Bookings.” *Journal of Revenue and Pricing Management* 7:2 (2008), 153–171.

- [Salganik et al. 06] M. Salganik, P. Dodds, and D. Watts. “Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market.” *Science* 311 (2006), 651.
- [Schelling 78] T. Schelling. *Micromotives and Macrobehaviour*. New York: Norton, 1978.
- [Sorensen 06] A. Sorensen. “Social Learning and Health Plan Choice.” *RAND Journal of Economics* 37:4 (2006), 929–945.
- [Vulcano et al. 10] G. Vulcano, G. van Ryzin, and W. Chahr. “On Practice–Choice–Based Revenue Management: An Empirical Study of Estimation and Optimization.” *Manufacturing and Service Operations Management* 12:3 (2010), 371–392.
- [Wierenga 08] B. Wierenga. *Handbook of Marketing Decision Models*. New York: Springer, 2008.
- [Woittiez and Kapteyn 98] I. Woittiez and A. Kapteyn. “Social Interactions and Habit Formation in a Model of Female Labour Supply.” *Journal of Public Economics* 70:2 (1998), 185–205.
- [Young 98] H. Young. *Individual Strategy and Social Structure*. Princeton: Princeton University Press, 1998.

Simla Ceyhan, Stanford University, 475 Via Ortega, Stanford, CA 94305, USA
(simlac@stanford.edu)

Mohammad Mousavi, Stanford University, 475 Via Ortega, Stanford, CA 94305, USA
(mousavi@stanford.edu)

Amin Saberi, Stanford University, 475 Via Ortega, Stanford, CA 94305, USA
(saberi@stanford.edu)

Received April 26, 2010; accepted February 22, 2011.