MARKET SHARE INDICATES QUALITY

AMIR BAN AND NATI LINIAL

Abstract. Market share and quality, or customer satisfaction, go hand in hand. Yet it is hard to find in the literature a clear formal statement to the effect that higher market share indicates higher quality. Indeed, such an inference would need detailed information about customer behavior. Moreover, even when that data is available, the validity of the inference is cast in doubt by common modes of behavior such as herding, the tendency to consume products due to their known popularity, or elitism, the opposite behavoir where customers associate mass popularity with lower quality. We investigate a model where customers are informed about their history with products and about global market share data. We find that it is in fact correct to make a Bayesian inference that the product with the higher market share has the better quality under few and rather unrestrictive assumptions on customer behavior.

1. Introduction

Common wisdom holds that a full restaurant is a good one, or certainly better than its empty neighbor. The purpose of this paper is to discover some minimal assumptions on the rationaly of customers under which this folk wisdom can be mathematically justified. For example, this conclusion certainly does not hold in places where the (admittedly strange) general preference is for food of poor taste. We formulate a simple model of a market in which customers have several products available to them. Each product has an innate unknown *quality* which is the probability that a customer who consumes it is *satisfied*. We find very mild sufficient conditions under which a larger market share indicates higher product quality.

Our approach is to model customer behavior with the least possible restrictions on customers' behavior. A large body of economical research indicates that customers' decisions use *bounded* rationality. Our approach also has the advantage of being independent of any particular model of benefit-maximizing strategy.

Intuitively, quality goes hand in hand with market share, and indeed a manufacturer's pursuit of quality is usually rationalized as a way of maximizing economic benefit. Empirical studies do not show conclusive evidence, though they (e.g. Anderson et. al. [1], Rust and Zahorik [5]) generally support a positive correlation between quality and market share.

We model quality as a probability for customer satisfaction. We consider quality to be a hidden, constant attribute of a product, which may be inferred, but not directly observed or learned from an authority. This is a widespread scenario that

Amir Ban amirban@netvision.net.il Center for the Study of Rationality, Hebrew University, Jerusalem, Israel.
Nati Linial nati@cs.huji.ac.il School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel.

customers realistically face. In the exceptions, e.g. when offered a Ferrari and a Fiat, price or other differentiation will typically also exist.

We ascribe to each customer a strategy of whether to consume each of the products, guided by market share and by one's own past personal history. When described as a behavioral strategy, i.e. by probabilities for consuming the product for each of the customer's information sets, we call it the customer's *partiality* strategy for that product. No connection between consumption of different products is assumed. E.g., a customer may consume all products simultaneously or none. We make no assumption that customers have uniform strategies, or that any customer's strategy is optimal.

We make two mild assumptions on customers' strategies, that customers (i) do not prefer negative over positive experiences with products, and (ii) do not prefer products with low market share over those with a higher one. Additionally, we assume an undifferentiated market, where products are *a priori* equal in the eyes of customers. Under these assumptions we can show the validity of inferences from market share to quality.

1.1. **Strategy Dependence on Personal History.** When customers (illogically) prefer dissatisfaction over satisfaction, market share clearly does not indicate quality. There are multiple ways to exclude or limit this from our framework. For example, we could require that customers' strategies be consistent with their *average* satisfaction with products. But this is already restrictive: It is, for example, not unreasonable to prefer a product used satisfactorily 18 times out of 20 trials over a product used just once satisfactorily; Or to give more weight to more recent trials.

Therefore we adopt a tamer restriction, which we call *monotonicity*: Namely, that customers recall the outcomes of their experiences with products, in the order that they happened, and if that history is definitely superior, on an experience-by-experience basis, their partiality strategy[1] to the product will be equal or higher. As an example, if a customer has a *fail-success-fail* history with a product (on the 3 occasions she elected to use it), then her partiality strategy after such a history will not be higher than if her history would have been *fail-success-success*, as the latter is superior by having a success where the former has failure, and is otherwise the same. No restriction is made on strategy after, e.g., the history *success-fail-success* (incomparable on an experience-by-experience basis), or *success-success* (incomparable due to a different number of experiences). Nor does it restrict the customer's strategy to other products, or other customers' strategies, as each can be formed independently within our framework.

This is possibly the lightest restriction on customer strategies we could make that conforms with common sense. When customers' strategies are guided solely by their history, we show it is sufficient to establish our result. Even for this restricted scenario, the conclusion is deeper than suggested by the assertion's simplicity: Attempted proofs must deal with a side result to which we allude in the Discussion: When quality varies with time, inferences from quality to market share or vice versa are, as a rule, invalid.

[1] Here and hereafter we use "partiality strategy" as shorthand for the probability of choosing the action of consumption under that strategy.

1.2. **Strategy Dependence on Market Share.** Customers may base their strategy on market share itself. This may take several forms: Customers may be fully or partially informed of market share, by, e.g. knowing product sales figures, or the ranking of the top-selling products. Smallwood and Conlisk[6] considered a market that evolves based on products having an intrinsic probability for breakdown and customers switching products randomly weighed by a function of market share. The present authors [2] considered a system where customers are influenced by history and reputation, where "reputation" under a suitable choice of model parameters represents market share. Word-of-mouth, i.e. asking or following others, is in effect a sampling of market share. Ellison and Fudenberg[4] considered a model of learning involving both personal history and word-of-mouth communication in which technologies perform stochastically based on an underlying quality parameter. Information cascades, starting with Bikhchandani et. al.[3], consider the inferences that observers can make on the quality of a service based on the customers queueing for that service and how informed those customers are known to be. They show this leads to *herding*, the phenomenon where customers accumulate due to the presence of others. Smallwood and Conlisk[6] as well as Ban and Linial[2] also show that lower quality products can maintain higher market share indefinitely.

Another potential feature of customer behavior we call *elitism*: Customers who intentionally avoid the most popular products. This may be due to a wish to differentiate oneself from the crowd, or to a belief that popular choices are second-rate, or any other reason.

Herding and elitism seem to cast doubt on our thesis. Herding, in particular, seems to pull the rug from underneath our sought conclusion. For example, if more than half of the customers at any point in time consume the market-leading product, and it alone, then market leadership is self-perpetuating regardless of product qualities and regardless of how other customers behave. No monotonicity assumptions are violated (for added credibility, assume leader-following customers are one-shot with no experiences to rely on), yet market share indicates nothing regarding product qualities. However, as we demonstrate, herding poses no problem to a market-to-quality inference: While a lower-quality product may sometimes prevail in market share, this will always have a lower probability than the alternative, and market share data *per se* is of no help in recognizing that such an anomaly is occurring. Defining a customer to be *weakly herding* if greater market share makes her more likely to consume a product, or has no effect on her behavior, we demonstrate that when all customers are weakly herding (in addition to being monotone on their product histories), market share is a valid signal for quality.

As for elitism, we believe, but do not analyze in the current paper, that if outweighed (in some sense) by herding, our thesis is still valid. Markets in which customer elitism is dominant turn out to be chaotic and difficult to analyze. In the Discussion we give an example where such a market does not adhere to our thesis. However, such markets seem far-fetched and so of low economic significance.

1.3. **No Other Differentiation.** Our result applies to undifferentiated markets, where all products are *a priori* equal in the eyes of customers. When customers distinguish between products by price, brand name, etc., or in captive markets, market share may be a reflection of the existing differentiation rather than of quality.

In our model, this translates to a requirement of *anonymity* of products in customer strategies, meaning that customers' strategies are invariant under a change of product labels.

In differentiated markets we are nevertheless able to state and prove a partial but significant result, namely, that the likelihood of market share leadership increases with quality.

1.4. **Organization of this paper.** The rest of this paper is organized as follows: Section 2 and 3 analyze markets where customers are guided entirely by their product history, with section 2 devoted to describing the model and section 3 stating and proving our proposition in such markets. Subsequently we analyze markets where customers are aware of product market share and take it into account, with section 4 devoted to refining the model for such markets, while section 5 states and proves our proposition. Formally speaking the results in Sections 4 and 5 subsume those of Sections 2 and 3, but we feel that this organization of the material makes it easier for the reader to follow. In Section 6 we state and prove an auxiliary theorem that is extensively used in our proofs. Conclusions are given in section 7.

2. Basic Model, When Only History Matters

In our model, customers make decisions regarding products in *rounds* of discrete time $t = 1, 2, \ldots$. At each round, a customer has an action set $\{C, N\}$, where $C :=$ consume the product, $N :=$ do not consume the product. If she consumes the product, she will, with probability given by the product's quality $q \in [0, 1]$, be satisfied, in which case the round is called an *S-round*, or else dissatisfied, in which case the round is called an *F-round*. If she chooses not to consume the product, the round is called an *N-round*.

A customer's *t-deep history* with a product, is a member of $\mathcal{H}_t := \{S, N, F\}^t$. Histories of depth up to t are denoted $\mathcal{H}_{\le t} := \bigcup_{k=0}^{t} \mathcal{H}_k$, and the set of all histories by $\mathcal{H} := \mathcal{H}_{\le\infty}$. For $Z \in \mathcal{H}_t$ we mark Z's *depth* $|Z| := t$. We denote by $\mathcal{H}^*_{\le t}$ the set of histories of depth up to t that end in a consumption event.

The customer's *partiality strategy*, $\sigma : \mathcal{H} \to [0, 1]$, is her behavioral strategy given her information set, which in this basic model is her history with the product at the time of decision. The partiality strategy is completely specified by specifying $\sigma(Z)$, the probability for action C, for each history $Z \in \mathcal{H}$.

ZV stands for $(Z(1), \ldots, Z(t), V) \in \mathcal{H}_{t+1}$, with $Z(k)$ standing for the event in round k.

The k'th tail of history Z, denoted $r^k(Z)$, is the history $(Z(1), \ldots, Z(t-k)) \in \mathcal{H}_{t-k}$ provided $t \ge k$ and the empty history otherwise. We use the shorthand $r(Z)$ for $r^1(Z)$. If Y is a tail of Z, Z is said to be an *extension* of Y.

We further define $S(Z)$ (resp. $F(Z)$, $N(Z)$) as the number of S-rounds (resp. F, N-rounds) in Z, i.e., the number of indices i for which $Z(i) = S$ (resp. F, N). The *consumption* of Z is defined as $con(Z) := S(Z) + F(Z)$. The *digest* of Z, denoted $dig(Z) \in \mathcal{H}_{con(Z)}$, is defined as the history that we obtain when we omit all the N-rounds from Z while maintaining the order of the remaining rounds.

For $Z \in \mathcal{H}$, we define the *implementation set* of Z which is denoted by $\Phi(Z)$. This is the set of all $Y \in \mathcal{H}$ such that $dig(Y) = dig(Z)$.

Let $Z_1, Z_2 \in \mathcal{H}_t$, with $con(Z_1) = con(Z_2)$, and let $D_1 := dig(Z_1), D_2 := dig(Z_2)$. We say that Z_1 is *superior* to Z_2, denoted $Z_1 \succeq Z_2$ if there is no index i for which $D_1(i) = F$ and $D_2(i) = S$.

A partiality strategy $\sigma(\cdot)$ is called *monotone* if

$$\sigma(Z_1) \geq \sigma(Z_2) \text{ whenever } Z_1 \succeq Z_2. \tag{2.1}$$

3. The Main Theorem When Only Product History Matters

In the current section we focus on the situation of a monotone partiality strategy that depends only on history. What can be said about the probability that the consumption up to time t, is $\geq x$ for arbitrary t and x? As the following theorem shows, this probability is a non-decreasing function of the product quality q.

Theorem 1. *Fix a monotone partiality strategy σ, and nonnegative integers t, x. Then*

$$\frac{d}{dq}\mathbf{P}\Big[con(Z) \geq x | Z \in \mathcal{H}_t\Big] \geq 0 \tag{3.1}$$

where the probability space is $\mathcal{H}$.

3.1. **Proof of Theorem 1.** We define a Markov chain on histories, i.e. a Markov chain with state space $\mathcal{H}$ that describes the possible transitions between histories and their probabilities. All transitions are from a member $Z \in \mathcal{H}_t$ to an extension $Z' \in \mathcal{H}_{t+1}$ with the following probabilities

$$\begin{array}{ll} Z \to ZS & \text{with probability } q \cdot \sigma(Z) \\ Z \to ZN & \text{with probability } 1 - \sigma(Z) \\ Z \to ZF & \text{with probability } (1-q) \cdot \sigma(Z) \end{array} \tag{3.2}$$

Consider the probability of reaching $Z \in \mathcal{H}$ as we start from the empty history and move along the Markov chain. It is convenient to express this probability as

$$c(Z) q^{S(Z)} (1-q)^{F(Z)}.$$

We refer to $c(Z)$ as the *ex-ante* function corresponding to strategy $\sigma(\cdot)$. Following from (3.2), its value is recursively defined by:

$$c(Z) = \begin{cases} 1 & Z = \emptyset \\ \sigma(r(Z))c(r(Z)) & Z(|Z|) \in \{S, F\} \\ \big[1 - \sigma(r(Z))\big]c(r(Z)) & Z(|Z|) = N \end{cases} \tag{3.3}$$

For example $c(FNSSN) = \sigma(\emptyset)[1-\sigma(F)]\sigma(FN)\sigma(FNS)[1-\sigma(FNSS)]$ where $\emptyset$ denotes the empty history. Observe that $c(Z)$ is a product of $|Z|$ factors. The factor has the form $\sigma(\cdot)$ for each consumption event, and $1-\sigma(\cdot)$ where the history has an N-event. The arguments of σ in the factors run over all $|Z|$ tails of Z.

It will be convenient to use an alternative description of the event in (3.1). Note that the consumption after t rounds is *at least* x iff the consumption is *exactly* x after *at most* t rounds. Furthermore, the shortest history with consumption x ends in a consumption event, i.e. is in $\mathcal{H}^*_{\leq t}$, and all such histories in $\mathcal{H}^*_{\leq t}$ are mutually disjoint events. In other words, the theorem can be equivalently stated as:

$$\frac{d}{dq}\mathbf{P}\Big[con(Z) = x | Z \in \mathcal{H}^*_{\leq t}\Big] \geq 0 \tag{3.4}$$

Let us try and provide some intuition for our next lemma. It states that the sum of the ex-ante function $c(\cdot)$ over the implementation set of a given digest *grows* as that digest is *improved* (in the sense defined above for the monotonicity property), provided the customer adheres to monotonicity. A customer's history is a process where both the customer and nature make random choices. The customer decides what to consume and her random choices are controlled by her partiality strategy. Nature chooses between success and failure, with randomness being controlled by the product quality. A history digest thus lists nature's choices, while the ex-ante function represents the total probability of the customer's choices, independently of nature. A digest's implementation set is therefore all histories with nature's choices given in advance.

By adopting this imaginary predetermination of nature's choices, we understand the *implementation depth* of a digest as the answer to the following question: Given that (unknown to the customer), nature's sequence of choices on product satisfaction is given, how many rounds will it take the customer to "get through it"? As shown in Lemma 1, and even more explicitly in Lemma 2, as this sequence of choices improves (with failures being replaced by successes), a customer observing monotonicity will "implement" the improved sequence more quickly.

The later parts of our proof show how this key property leads to a proof of the entire theorem.

Lemma 1. *Let $D_1, D_2 \in \{S, F\}^l$ be two histories satisfying $D_1 \succeq D_2$. Let the partiality strategy $\sigma(\cdot)$ be monotone, with $c(\cdot)$ the corresponding ex ante function. Then for all positive t:*

$$\sum_{Z \in \mathcal{H}^*_{\leq t} \cap \Phi(D_1)} c(Z) \geq \sum_{Z \in \mathcal{H}^*_{\leq t} \cap \Phi(D_2)} c(Z) \tag{3.5}$$

Proof. At the outset of the proof, we wish to justify its complexities, which may perhaps suprise the reader, who will question their necessity. The problem stems from the fact that the $c(\cdot)$ function is *not* monotone, even if the partiality strategy $\sigma(\cdot)$ is, i.e. it is not true that $c(Z_1) \geq c(Z_2)$ whenever $Z_1 \succeq Z_2$. The reason is the presence of $1 - \sigma(\cdot)$ factors in the definition of the ex-ante function (see (3.3)) which decrease, rather than increase when the history argument is improved. Indeed, were the ex-ante function monotone, the lemma could be trivially proved by term-by-term comparison. However, this is not the case, and rather than being trivial, the lemma states an inequality which is algebraically difficult to prove even in simplified special cases.[2]

The non-monotonicity of the ex-ante function is key to understanding why not only the lemma, but the main result is an algebraically advanced proposition. To settle its veracity, we rely on a reduction of the partiality strategy to integer values.

[2]For example, for $t = 3, D_1 = SS, D_2 = FS$, mark:

$$0 \leq a := \sigma(\emptyset) \leq 1$$
$$0 \leq b' := \sigma(F) \leq b := \sigma(S) \leq 1$$
$$0 \leq c := \sigma(N) \leq 1$$
$$0 \leq d' := \sigma(FS) \leq d := \sigma(SS) \leq 1$$
$$0 \leq e' := \sigma(FN) \leq e := \sigma(SN) \leq 1$$
$$0 \leq f' := \sigma(NF) \leq f := \sigma(NS) \leq 1$$

The reduction, and the proof of its validity, is detailed in a separate section, Section 6.

Let $m = |\mathcal{H}_{\leq t}|$, and let us associate the coordinates of $\mathbb{R}^m$ with members of $\mathcal{H}_{\leq t}$, with each coordinate value identified with the strategy value $\sigma(\cdot)$ for that member. We define a polytope $P \subseteq \mathbb{R}^m$ by the inequalities that determine monotone partiality strategies:

$$\forall i \in [m] \qquad 0 \leq \sigma(Z_i) \leq 1 \tag{3.6}$$

$$\forall i, j \in [m], i \neq j, Z_i \succeq Z_j \qquad \sigma(Z_i) \geq \sigma(Z_j) \tag{3.7}$$

Every monotone partiality strategy σ corresponds to a point in P, whose coordinates are σ's values for each history in $\mathcal{H}_{\leq t}$. (Strategies for histories outside $\mathcal{H}_{\leq t}$ are irrelevant to this lemma). Define $F(\sigma)$ to be:

$$F(\sigma) := \sum_{Z \in \mathcal{H}^*_{\leq t} \cap \Phi(D_1)} c(Z) - \sum_{Z \in \mathcal{H}^*_{\leq t} \cap \Phi(D_2)} c(Z) \tag{3.8}$$

Then the lemma asserts:

$$\min_{\sigma \in P} F(\sigma) \geq 0 \tag{3.9}$$

We can apply Theorem 6 (Section 6) to our situation and conclude that it suffices to prove our lemma for *pure* partiality strategies, that take only the values 0 or 1. Identify each coordinate x_i in the theorem with $\sigma(Z_i)$ here. The conclusion holds, since, as we observe below, the function F satisfies the condition of the theorem.

By definition (3.3) $c(\cdot)$ is multilinear, being a product of up to t linear factors, each involving $\sigma(Z)$ for some $Z \in \mathcal{H}_{\leq t}$. Therefore, by (3.8), F is multilinear, as needed. We next show that F satisfies the condition that is required in Theorem 6. Each factor in the expression for $c(\cdot)$ involves a history of different depth. On the other hand the monotonicity inequalities (3.7) compare only between same-depth histories. Thus the different factors belong to different components of G, as defined in Theorem 6.

It remains to prove the lemma for pure partiality strategies. This we now set to do. When partiality strategy $\sigma(\cdot)$ is restricted to 0 or 1 values, then necessarily so is the corresponding ex-ante function $c(\cdot)$. The lemma's assertion, restated for this case, is:

Given $D_1 \succeq D_2$:

$$|\{Z | Z \in \mathcal{H}^*_{\leq t} \cap \Phi(D_1), c(Z) \neq 0\}| \geq |\{Z | Z \in \mathcal{H}^*_{\leq t} \cap \Phi(D_2), c(Z) \neq 0\}| \tag{3.10}$$

To prove this, we first show that both sides of (3.10) are at most 1. For suppose that two distinct histories Z', Z'' contribute to the left-hand side. Since the digests of the histories are the same, anywhere they differ one must have either an S or an F round where the other has an N round. Consider the earliest such difference: Since partiality strategy values are restricted to 0 and 1 values, one of the histories has zero probability for that round, and so it is impossible for both $c(Z')$ and $c(Z'')$ to be non-zero. Similarly the right-hand side cannot have two distinct histories.

The implementations of D_1 are SSN, SNS and NSS and of D_2, FSN, FNS and NFS. The lemma claims the true but non-obvious:

$$ab(1-d) + a(1-b)e + (1-a)cf \geq ab'(1-d') + a(1-b')e' + (1-a)cf'$$

Another way of reaching the same conclusion is to observe that, with partiality strategies set to either 0 or 1, and with history digests set in advance, the Markov chain for histories is deterministic. Under these deterministic circumstances, the minimum depth at which a given digest is implemented is called the digest's *implementation depth* (which may be infinite). The number of histories contributing to the left (right) side of (3.10) is 0 or 1 depending on whether the implementation depth of D_1 (D_2) is $> t$ or $\leq t$, respectively.

Therefore it is enough to prove that when the implementation depth of D_2 is $\leq t$, so is the implementation depth of D_1. The following sub-lemma settles this point:

Lemma 2. *Let $D_1 \succeq D_2$ be two l-deep histories, $D_1, D_2 \in \{S, F\}^l$. Let the partiality strategy $\sigma(\cdot)$ be monotone and pure, and let $c(\cdot)$ be the ex ante function based on it. If there exists $Z_2 \in \Phi(D_2)$ with depth t such that $c(Z_2) = 1$, then there also exists $Z_1 \in \Phi(D_1)$ with depth $\leq t$ such that $c(Z_1) = 1$.*

Proof. By induction on the history's depth t. When $t = 0$ the assertion is vacuously true. Assume the lemma proven for histories of depth up to $t-1$. Let Z_2 be a t-deep history with digest D_2 and $c(Z_2) = 1$.

Assume that the lemma is not true and the implementation depth of D_1 is greater than t. If $Z_2(t) = N$, then $r(Z_2)$ has the same digest as Z_2 and depth $t-1$, so invoking the induction hypothesis causes a contradiction. Otherwise, as Z_2 ends in a consumption event then (i) $\sigma(r(Z_2)) = 1$ (ii) $dig(r(Z_2)) = r(D_2)$. Since $r(Z_2)$ is $(t-1)$-deep and $r(D_1) \succeq r(D_2)$, by the induction hypothesis there is an implementation of $r(D_1)$ in depth $\leq t-1$. The only way for there *not* to be an implementation of D_1 in depth $\leq t$ is if $\sigma(r(Z_1)) = 0$. But $r(Z_1) \succeq r(Z_2)$ so $\sigma(r(Z_1)) < \sigma(r(Z_2))$ violates monotonicity: A contradiction. □

This proves (3.10) and the lemma for pure partiality strategies, which, as already demonstrated, is sufficient to prove the lemma. □

Lemma 3. *Let u, v be positive integers. Let the partiality strategy $\sigma(\cdot)$ be monotone, and let $c(\cdot)$ be the ex ante function based on it. Then for all positive t:*

$$(u+1) \sum_{\substack{Z \in \mathcal{H}^*_{\leq t} \\ S(Z)=u+1 \\ F(Z)=v}} c(Z) \geq (v+1) \sum_{\substack{Z \in \mathcal{H}^*_{\leq t} \\ S(Z)=u \\ F(Z)=v+1}} c(Z) \tag{3.11}$$

Proof. Let Y be a history digest with $S(Y) = u+1$ and $F(Y) = v$. Let Y' be derived from Y by altering one S round in Y to F, so that $S(Y') = u$ and $F(Y') = v+1$. Moreover, $Y \succeq Y'$. By lemma 1:

$$\sum_{Z \in \mathcal{H}^*_{\leq t} \cap \Phi(Y)} c(Z) \geq \sum_{Z \in \mathcal{H}^*_{\leq t} \cap \Phi(Y')} c(Z) \tag{3.12}$$

Sum this inequality over all such possible pairs $(Y, Y') \in \mathcal{H}^*_{\leq t} \times \mathcal{H}^*_{\leq t}$. On the left-hand side each history $Z \in \mathcal{H}^*_{\leq t}$ with $S(Z) = u+1, F(Z) = v$ appears $u+1$ times, as each has $u+1$ S-rounds that can be changed to F. On the right-hand side each history $Y' \in \mathcal{H}^*_{\leq t}$ with $S(Y') = u, F(Y') = v+1$ appears $v+1$ times, as each has $v+1$ F-rounds that could be changed from S. Hence the assertion of the lemma is this sum-total inequality. □

We now proceed to prove the theorem:

Proof. Multiply both sides of (3.11) by $q^u(1-q)^v$. This is a positive quantity as $0 \le q \le 1$. We get:

$$(u+1) \sum_{\substack{Z \in \mathcal{H}^*_{\le t} \\ S(Z)=u+1 \\ F(Z)=v}} c(Z) q^u (1-q)^v \ge (v+1) \sum_{\substack{Z \in \mathcal{H}^*_{\le t} \\ S(Z)=u \\ F(Z)=v+1}} c(Z) q^u (1-q)^v \tag{3.13}$$

Equivalently:

$$\sum_{\substack{Z \in \mathcal{H}^*_{\le t} \\ S(Z)=u+1 \\ F(Z)=v}} c(Z) S(Z) q^{S(Z)-1} (1-q)^{F(Z)} \ge \sum_{\substack{Z \in \mathcal{H}^*_{\le t} \\ S(Z)=u \\ F(Z)=v+1}} c(Z) F(Z) q^{S(Z)} (1-q)^{F(Z)-1} \tag{3.14}$$

Given a positive integer x, sum (3.14) over all pairs of positive integers u, v with $u + v = x - 1$. This yields:

$$\sum_{\substack{Z \in \mathcal{H}^*_{\le t} \\ con(Z)=x}} c(Z) S(Z) q^{S(Z)-1} (1-q)^{F(Z)} \ge \sum_{\substack{Z \in \mathcal{H}^*_{\le t} \\ con(Z)=x}} c(Z) F(Z) q^{S(Z)} (1-q)^{F(Z)-1}$$

Or:

$$\sum_{\substack{Z \in \mathcal{H}^*_{\le t} \\ con(Z)=x}} c(Z) \Big[S(Z) q^{S(Z)-1} (1-q)^{F(Z)} - F(Z) q^{S(Z)} (1-q)^{F(Z)-1} \Big] \ge 0$$

$$\Rightarrow \sum_{\substack{Z \in \mathcal{H}^*_{\le t} \\ con(Z)=x}} c(Z) \frac{d}{dq} \Big[q^{S(Z)} (1-q)^{F(Z)} \Big] \ge 0$$

$$\Rightarrow \frac{d}{dq} \mathbf{P} \Big[con(Z) = x | Z \in \mathcal{H}^*_{\le t} \Big] \ge 0$$

as claimed. □

It follows from Theorem 1 that the expected consumption is non-decreasing in the quality:

Corollary 1. *If the partiality strategy is monotone, then for any time t:*

$$\frac{d}{dq} \mathbb{E} \Big[con(Z) | Z \in \mathcal{H}_t \Big] \ge 0 \tag{3.15}$$

where the probability space is $\mathcal{H}$.

Proof. Since:

$$\mathbb{E} \Big[con(Z) | Z \in \mathcal{H}_t \Big] = \sum_{x=1}^{\infty} \mathbf{P} \Big[con(Z) \ge x | Z \in \mathcal{H}_t \Big] \tag{3.16}$$

This follows from Theorem 1. □

Having proved Theorem 1 for the consumption of a single customer, we state and prove an equivalent theorem for an entire market, i.e. that, provided each customer's partiality strategy is monotone, a product's market share (i.e. its total consumption) stochastically dominates the market share of a product of lesser quality:

Theorem 2. *Let each of n customers have a monotone partiality strategy. Given a time t and an integer x:*

$$\frac{d}{dq}\,\mathbf{P}\Big[\sum_{j=1}^{n} con(Z_j) \geq x | \forall j \in [n], Z_j \in \mathcal{H}_t\Big] \geq 0 \tag{3.17}$$

where the probability space is $\mathcal{H}^n$.

Proof. The proof proceeds by induction on n. For $n = 1$ this is just Theorem 1. Assume the statement true for up to $n-1$ customers. Mark $a_n(x) := \mathbf{P}\Big[\sum_{j=1}^{n} con(Z_j) \geq x | \forall j \in [n], Z_j \in \mathcal{H}_t\Big]$ and $b(y) := \mathbf{P}\Big[Z_n \geq y | Z_n \in \mathcal{H}_t\Big]$. Then:

$$\begin{aligned}
\frac{d}{dq}a_n(x) &= \frac{d}{dq}\sum_{y=-\infty}^{\infty} a_{n-1}(x-y)\,\mathbf{P}\Big[Z_n = y | Z_n \in \mathcal{H}_t\Big] = \\
&= \frac{d}{dq}\sum_{y=-\infty}^{\infty} a_{n-1}(x-y)b(y) - \frac{d}{dq}\sum_{y=-\infty}^{\infty} a_{n-1}(x-y)b(y+1) = \\
&= \sum_{y=-\infty}^{\infty} \frac{d}{dq}a_{n-1}(x-y)b(y) + \sum_{y=-\infty}^{\infty} a_{n-1}(x-y)\frac{d}{dq}b(y) - \\
&- \sum_{y=-\infty}^{\infty} \frac{d}{dq}a_{n-1}(x-y)b(y+1) - \sum_{y=-\infty}^{\infty} a_{n-1}(x-y)\frac{d}{dq}b(y+1)
\end{aligned}$$

Changing variables in the last term $y+1 \to y$ and combining, this results in:

$$\begin{aligned}
&\sum_{y=-\infty}^{\infty} \frac{d}{dq}a_{n-1}(x-y)\,\mathbf{P}\Big[Z_n = y | Z_n \in \mathcal{H}_t\Big] + \\
&\sum_{y=-\infty}^{\infty} \mathbf{P}\Big[\sum_{j=1}^{n-1} con(Z_j) = x-y | \forall j \in [n-1], Z_j \in \mathcal{H}_t\Big]\frac{d}{dq}b(y) \geq 0
\end{aligned}$$

since all factors in the above expression are non-negative. □

The main result can now be stated and proved: If two products are interchangeable in the eyes of the customers, and if there is no prior cause to believe that one of the products has the better quality, then from the observation of a higher market share for one of the products one can infer that it has the better quality.

Theorem 3. *Let the partiality strategies of all customers for products 1 and 2 be monotone, and let each customer's strategy for product 1 be the same as for product 2. Let products 1, 2 have possibly different qualities q_1, q_2 respectively, with symmetric prior. Let the history of customer $j \in [n]$ with product $i \in [m]$ be*

$Z_{ij} \in \mathcal{H}_t$. Then:

$$\mathbf{P}\Big[q_1 \geq q_2 | \sum_{j=1}^{n} con(Z_{1j}) > \sum_{j=1}^{n} con(Z_{2j})\Big] \geq \mathbf{P}\Big[q_2 \geq q_1 | \sum_{j=1}^{n} con(Z_{1j}) > \sum_{j=1}^{n} con(Z_{2j})\Big]$$

where the probability space is $\mathcal{H}^{mn}$.

Proof. Mark $\omega_1 := \sum_{j=1}^{n} con(Z_{1j})$ and $\omega_2 := \sum_{j=1}^{n} con(Z_{2j})$.

As the products are interchangeable $\mathbf{P}[\omega_1 > \omega_2 | q_1 = q_2] = \mathbf{P}[\omega_2 > \omega_1 | q_1 = q_2]$. By theorem 2:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_1 > \omega_2 | q_1 = q_2] \tag{3.18}$$

$$\mathbf{P}[\omega_2 > \omega_1 | q_1 \geq q_2] \leq \mathbf{P}[\omega_2 > \omega_1 | q_1 = q_2] \tag{3.19}$$

Therefore:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_2 > \omega_1 | q_1 \geq q_2] \tag{3.20}$$

As the products are interchangeable, $\mathbf{P}[\omega_2 > \omega_1 | q_1 \geq q_2] = \mathbf{P}[\omega_1 > \omega_2 | q_2 \geq q_1]$, therefore:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_1 > \omega_2 | q_2 \geq q_1] \tag{3.21}$$

¿From which the theorem follows by Bayes' theorem and the symmetric prior on q_1, q_2. □

4. Model with Market Share Observations

We now generalize our model to the case where the market share of the products is known to customers. Customers can base their partiality strategies on market share information, as well as on their individual history with the products. We need to define market share:

Let there be n customers and m products. Let $Z_{ij} \in \mathcal{H}_t$ be customer j's t-deep history with product i. We define a (t-deep) *history ensemble* Z as a set of histories for each customer-product combination $Z := \{Z_{ij} \in \mathcal{H}_t, \forall i \in [m], j \in [n]\}$.

The set of t-deep history ensembles is denoted by $\mathcal{G}_t$.

The τ-deep ($0 \leq \tau \leq t$) tail of a history ensemble Z is defined as $r^{t-\tau}(Z) := \{r^{t-\tau}(Z_{ij}) \in \mathcal{H}_\tau, \forall i \in [m], j \in [n]\}$.

Given a t-deep history ensemble Z, and an initial market share $A := (A_1, \ldots, A_m)$, the market share of product $i \in [m]$ after round τ is the total number of units consumed of product m up to round τ, and is denoted by $\Omega_i(Z, A, \tau)$:

$$\Omega_i(Z, A, \tau) = A_i + \sum_{j=1}^{n} con(r^{t-\tau}(Z_{ij})) \tag{4.1}$$

All customers are aware of the round-τ market share of all products[3] when they make their consumption decisions at round $\tau + 1$.

$\Omega(Z, A, \tau)$ denotes the vector of all product market shares $(\Omega_1(Z, A, \tau), \ldots, \Omega_m(Z, A, \tau))$.

The initial market share vector $A \equiv \Omega(Z, A, 0)$ is the market share vector before round 1, and its value is extraneous to the model.

[3]See the Discussion for some comments on how things behave when customers may have more information about past market shares.

The partiality strategy of customer $j \in [n]$ to product $i \in [m]$ after round t, $\sigma_{ij}(Z_{ij}, \Omega(Z, A, t))$, is defined as the probability that the customer will consume product i at round $t+1$. It is a behavioral strategy that depends on the customer's information set which consists of Z_{ij}, the customer's history with product i, and the market share $\Omega(Z, A, t)$ known after round t.

We generalize the definition of a t-deep history ensemble to a t-deep *partial* history ensemble, where member histories do not necessarily have the same depth: $Z := \{Z_{ij} \in \mathcal{H}_{\leq t}, \forall i \in [m], j \in [n]\}$.

The set of t-deep partial history ensembles is denoted by $\mathcal{G}_{\leq t}$, and the set of all history ensembles by $\mathcal{G} := \mathcal{G}_{\leq \infty}$.

A t-deep partial history ensemble Z is said to occur if, at time t, all actual histories are extensions of their respective partial histories in Z. The set of actual t-deep histories that extend a t-deep partial history ensemble is denoted by $\Psi_t(Z)$:

$$\Psi_t(Z) := \{Y \in \mathcal{G}_t | \forall i \in [m], j \in [n], \exists t_{ij} \geq 0, r^{t_{ij}}(Y_{ij}) = Z_{ij}\} \tag{4.2}$$

We define the digest of a history ensemble to be the ensemble of digests of each of its histories: $D = dig(Z) \Rightarrow \forall i \in [m], j \in [n], D_{ij} = dig(Z_{ij})$.

Define:

$$S_i(Z) := \sum_{j=1}^{n} S(Z_{ij}) \tag{4.3}$$

$$F_i(Z) := \sum_{j=1}^{n} F(Z_{ij}) \tag{4.4}$$

$$Q_i(Z) := q_i^{S_i(Z)}(1 - q_i)^{F_i(Z)} \tag{4.5}$$

$$Q(Z) := \prod_{i=1}^{m} Q_i(Z) \tag{4.6}$$

We define monotonicity similarly to how we defined it in Section 2: A partiality strategy σ_{ij} is monotone if for every market share vector ω and for every history pair Z_1, Z_2 satisfying $Z_1 \succeq Z_2$, $\sigma_{ij}(Z_1, \omega) \geq \sigma_{ij}(Z_2, \omega)$.

We introduce a condition on customers' response to market data that we call *weak herding*. As we show, if weak herding holds and if customers are monotone, then a result similar to Theorem 3 holds. Namely, market share still indicates quality. Compared with the previous sections, customers are less restricted in forming their partiality strategies under these assumptions.

A customer j is called *weakly herding* if for every product i, time t, t-deep history Z_{ij} and market share vector $\omega = \{\omega_1, \ldots, \omega_m\}$, $\sigma_{ij}(Z_{ij}, \omega)$ is non-decreasing in ω_i. A customer j is called *competitively weakly herding* if for every product i, time t, t-deep history Z_{ij} and market share vector $\omega = \{\omega_1, \ldots, \omega_m\}$, $\sigma_{ij}(Z_{ij}, \omega)$ is non-decreasing in ω_i *and* non-increasing in ω_k for all $k \neq i$. In particular, a customer who, as in our basic model, is unaware of market share or disregards it, is both weakly herding and competitively weakly herding.

Weak herding is a natural response to market share data: The more a product has been consumed, the more a customer who is aware of that fact is disposed to consume it. Competitive weak herding makes it possible to base partiality strategies on a product's *order* in market share data, e.g. on whether or not a product is the market leader in consumption. We shall be able to prove our thesis for both kinds of responses, though for competitive weak herding we shall have to limit the number of products.

Additionally, we define *anonymity* for products, the property that products are *a priori* equal in the eyes of customers. For anonymous products, partiality strategies do not depend on the label of a product but only on its data. Formally, let $Z = ((Z_{11}, \ldots, Z_{1n}), (Z_{21}, \ldots, Z_{2n}), \ldots, (Z_{m1}, \ldots, Z_{mn}))$ and $\omega = (\omega_1, \omega_2, \ldots, \omega_m)$. Define the permutations $K_{12}(Z) := ((Z_{21}, \ldots, Z_{2n}), (Z_{11}, \ldots, Z_{1n}), \ldots, (Z_{m1}, \ldots, Z_{mn}))$ and $K_{12}(\omega) := (\omega_2, \omega_1, \ldots, \omega_m)$. Then products $1, 2$ are *anonymous* if, for each customer j, for each history Z and for each market share vector ω:

$$\sigma_{1j}(Z, \omega) = \sigma_{2j}(K_{12}(Z), K_{12}(\omega)) \tag{4.7}$$

5. Theorem with Market Share Observation

Theorem 4. (1) *Assume that all customers are monotone and weakly herding and fix some initial conditions $A = (A_1, \ldots, A_m)$. Then for all times t:*

$$\frac{d}{dq_1}\mathbf{P}\Big[\Omega_1(Z, A, t) > \Omega_2(Z, A, t) | Z \in \mathcal{G}_t\Big] \geq 0 \tag{5.1}$$

where $q_i \in [0, 1]$ is the quality of product i for $i = 1, \ldots, m$, and the probability space is $\mathcal{G}$.

(2) *The same holds when all customers are competitively weakly herding, rather than weakly herding, and there are two products ($m = 2$).*

Proof. Our proof is modeled on the proof of Theorem 1, but with significant changes. The proof proceeds in two steps: First, we show that it suffices to prove the theorem when all partiality strategies are pure (0/1 valued), and then prove it for pure partiality strategies.

Following (3.2), we note that, for every customer-product pair, histories form a Markov chain describing the possible transitions between histories and their probability. Further following (3.3), we define a customer-product specific ex-ante function $c_{ij}(Z, A)$ of customer i, product j. The definition is similar to that given by (3.3):

$$c_{ij}(Z, A) = \begin{cases} 1 & Z_{ij} = \emptyset \\ \sigma_{ij}(r(Z), \Omega(Z, A, t-1)) c_{ij}(r(Z), A) & Z_{ij}(t) \in \{S, F\} \\ \big[1 - \sigma_{ij}(r(Z), \Omega(Z, A, t-1))\big] c_{ij}(r(Z), A) & Z_{ij}(t) = N \end{cases} \tag{5.2}$$

where $t = |Z_{ij}|$.

The total ex-ante function $c(Z, A)$ is defined as the product of all customer-product specific ex-ante functions. The ex-ante probability of a history ensemble, i.e. the probability of its occurrence, is then calculated from the total ex-ante function and from the product qualities:

$$c(Z, A) = \prod_{i=1}^{m} \prod_{j=1}^{n} c_{ij}(Z, A) \tag{5.3}$$

$$\mathbf{P}[Z] = \mathbf{P}[\Psi_t(Z)] = c(Z, A) Q(Z) \tag{5.4}$$

The theorem, therefore, asserts that:

$$\frac{d}{dq_1} \sum_{\substack{Z \in \mathcal{G}_t \\ \Omega_1(Z,A,t) > \Omega_2(Z,A,t)}} c(Z,A)Q(Z) \geq 0 \tag{5.5}$$

Equivalently:

$$\sum_{\substack{Z \in \mathcal{G}_t \\ \Omega_1(Z,A,t) > \Omega_2(Z,A,t)}} c(Z,A)\left[\frac{S_1(Z)}{q_1} - \frac{F_1(Z)}{1-q_1}\right]Q(Z) \geq 0 \tag{5.6}$$

We now proceed to apply Theorem 6 (see Section 6): We consider a set of variables x consisting of all partiality strategy values of $\sigma_{ij}, \forall i \in [m], j \in [n]$, each for all possible histories in $\mathcal{H}_{\leq t}$ and for all possible market share vectors (with values that w.l.o.g. may be limited below some bound). We define $F(x)$ as the left-hand side of (5.6):

$$F(x) := \sum_{\substack{Z \in \mathcal{G}_t \\ \Omega_1(Z,A,t) > \Omega_2(Z,A,t)}} c(Z,A)\left[\frac{S_1(Z)}{q_1} - \frac{F_1(Z)}{1-q_1}\right]Q(Z) \tag{5.7}$$

Let us return to the notion of a partiality strategy. Given all the relevant data, i.e., history and market share information it outputs a probability (of a specific customer consuming the specific product in the next round). Let M denote the number of such possible data sets[4]. We consider the polytope $P \subseteq \mathbb{R}^M$ of all possible ensembles of (all customers-products') partiality strategies. Since strategy values are probabilities, $P \subseteq [0,1]^M$. Monotonicity and weak herding are expressed by linear inequalities as follows: For each history pair $Z, Z' \in \mathcal{H}_{\leq t}$, and for each market share vector pair ω, ω', and for each product $i \in [m]$ and each customer $j \in [n]$:

$$(Z \succeq Z') \wedge (\omega_i \geq \omega'_i) \Rightarrow \sigma_{ij}(Z,\omega) \geq \sigma_{ij}(Z',\omega') \tag{5.8}$$

While monotonicity and competitive weak herding are expressed by linear inequalities as follows: For each history pair $Z, Z' \in \mathcal{H}_{\leq t}$, and for each market share vector pair ω, ω', and for each product $i \in [m]$ and each customer $j \in [n]$:

$$(Z \succeq Z') \wedge (\omega_i \geq \omega'_i) \wedge (\forall k \neq i, \omega_k \leq \omega'_k) \Rightarrow \sigma_{ij}(Z,\omega) \geq \sigma_{ij}(Z',\omega') \tag{5.9}$$

This is a complete decription of polytope P, since the partiality strategies of different customers are independent of each other. The same is true for different products, as product anonymity has not been required in the current theorem.

The function $F(x)$ satisfies the conditions of Theorem 6. The ex-ante function $c(Z,A)$ is multilinear in x (its coefficients in the terms of $F(x)$ are scalar constants in the current context). Furthermore, note that (5.8) and (5.9) have inequalities only between histories of the same depth, as monotonicity is defined only for same-depth histories, and involves only same-customer, same-product partiality strategies. Since each customer-product specific ex-ante function has only one factor for each depth, each term of $F(x)$ depends on at most one member of each connected component of G, the graph with vertex set $[M]$ and edges defined by (5.8) (for weak herding) or by (5.9) (for competitive weak herding).

[4]Specifically, $M < mn3^{t+1}(2tn)^m$.

Consequently, $F(x)$ attains its minimum at some all-integer x. Therefore, it is enough to prove the theorem for pure (0/1-valued) partiality strategies.

This we now set to do. The theorem, stated in the form of (5.6), restricted to pure partiality strategies, reads as follows:

$$\sum_{\substack{Z\in\mathcal{G}_t\\ \Omega_1(Z,A,t)>\Omega_2(Z,A,t)\\ c(Z,A)=1}} \left[\frac{S_1(Z)}{q_1} - \frac{F_1(Z)}{1-q_1}\right] Q(Z) \geq 0 \tag{5.10}$$

Definition 1. *For $Z \in \mathcal{H}$, we denote by $\Phi^-(Z)$ the* contraction set *of Z. This is the set of all $Y \in \mathcal{H}$ such that $dig(Y) = r^m(dig(Z))$ for some $m \geq 0$.*

Note that the contraction set of a history is the union of the implementation sets of that history and of all its tails.

Here is an informal explanation of the following lemma. Consider a t-deep history ensemble in which each individual history has a t-deep implementation. If this history ensemble is improved (i.e. some F-events are changed to S-events) with respect to product 1, then in the improved history ensemble: (i) Implementation depths of all customer histories of product 1 are shortened or unchanged (i.e. are at most t) (ii) Implementation depths of all customer histories of products other than 1 are lengthened or unchanged (i.e. may be only partially implemented by time t).

We refer the reader to the preamble we gave to Lemma 1 in Section 3, and note that the current lemma is the analog of Lemma 2, in the current more complex setting: Fixing the digests in a history ensemble may be viewed as predetermining nature's choices in advance, and then asking how much time it will take the market (compared to a single customer in Section 3) to "get through it", under the assumption that all customers observe monotonicity and weak-herding.

Lemma 4. (1) *Let $Z' \in \mathcal{G}_t$ and let $D' = dig(Z')$ be the ensemble of all digests of histories in Z'. Assume $c(Z', A) = 1$. Let $D \in \mathcal{G}_{\leq t}$ be some other ensemble of digests satisfying $\forall j \in [n]$, $D_{1j} \succeq D'_{1j}$ and $\forall i \neq 1, D_{ij} = D'_{ij}$. Suppose that all customers are weakly herding, and all partiality strategies $\sigma_{ij}(\cdot)$, $\forall i \in [m], j \in [n]$ are monotone and pure. Let $c(\cdot)$ be the ex ante function corresponding to these strategies and to some given initial market share vector $A = (A_1, \ldots, A_m)$.*

Then:

(a) *There exists a t-deep partial history ensemble $Z \in \mathcal{G}_{\leq t}$ such that $c(Z, A) = 1$, where $\forall j \in [n], Z_{1j} \in \mathcal{H}_{\leq t} \cap \Phi(D_{1j})$, and $\forall j \in [n], i \neq 1, Z_{ij} \in \mathcal{H}_t \cap \Phi^-(D_{ij}))$.*

(b) *For every $Y \in \Psi_t(Z)$, product 1's market share in Y is at least as large as in Z', while for products other than 1, the market share in Y is at most as large as in Z'.*

(c)

$$\frac{Q(Z)}{Q_1(Z)} \geq \frac{Q(Z')}{Q_1(Z')} \tag{5.11}$$

(2) *The same holds when all customers are competitively weakly herding, rather than weakly herding, and there are two products ($m = 2$).*

Proof. We first note that part 1a of the lemma implies part 1b, since

$$\Omega_1(Y,A,t) \geq A_1 + \sum_{i=1}^{n} con(Z_{1i}) = A_1 + \sum_{i=1}^{n} con(Z'_{1i}) = \Omega_1(Z',A,t) \tag{5.12}$$

$$\forall i \neq 1 : \Omega_i(Y,A,t) = A_i + \sum_{i=1}^{n} con(Z_{ij}) \leq A_i + \sum_{i=1}^{n} con(Z'_{ij}) = \Omega_i(Z',A,t) \tag{5.13}$$

We prove part 1a (and hence also part 1b) by induction on t. For $t = 0$ the claim is vacuously true. Assume part 1a and part 1b proven for $t-1$, and we proceed to the induction step:

Consider the ensemble $r(Z') \in \mathcal{G}_{t-1}$. Its ensemble of digests $d' := dig(r(Z'))$ has a digest for each $i \in [m]$, $j \in [n]$ that is either (i) equal to D'_{ij} whenever the last round of the specific history was an N-round, i.e. when $Z'_{ij}(t) = N$, or (ii) equal to $r(D'_{ij})$ otherwise, i.e. when $Z'_{ij}(t) \in \{S, F\}$.

Construct the history ensemble d from D as follows: For each $i \in [m]$, $j \in [n]$, if $Z'_{ij}(t) = N$ set $d_{ij} = D_{ij}$, otherwise set $d_{ij} = r(D_{ij})$. Clearly $d \in \mathcal{G}_{\leq t-1}$, and we may apply the induction hypothesis on $r(Z')$, and on d, as d satisfies the requirements on D, namely, for each $j \in [n]$, $d_{1j} \succeq d'_{1j}$, and for each $i \neq 1$, $j \in [n]$, $d_{ij} = d'_{ij}$.

By part 1a there therefore exists $z \in \mathcal{G}_{\leq t-1}$ such that $c(z, A) = 1$, where $\forall j \in [n], z_{1j} \in \mathcal{H}_{\leq t-1} \cap \Phi(d_{1j})$, and $\forall j \in [n], i \neq 1, z_{ij} \in \mathcal{H}_{t-1} \cap \Phi^{-}(d_{ij}))$.

We shall prove that for every $j \in [n]$, D_{1j} has an implementation depth at most t. If $d_{1j} = D_{1j}$, then z_{1j} is such an implementation with depth $\leq t-1 < t$. Otherwise, i.e. if $Z'_{1j}(t) \neq N$, then as shown above $d_{1j} = r(D_{1j})$ has an implementation z_{1j} with depth $\leq t-1$. Furthermore, $Z'_{1j}(t) \neq N$ implies $\sigma_{1j}(r(Z'_{1j}), \Omega(r(Z'), A, t-1)) = 1$.

Since $r(D_{1j})$ has an implementation z_{1j} of depth at most $t-1$, the only way that D_{1j} will *not* have an implementation of depth $\leq t$ is for customer j to have zero partiality strategy for product 1 at all times starting at the depth of z_{1j} and up to and including $t-1$. But this is not possible: A customer who has zero partiality strategy up to time $t-1$ will necessarily have non-zero partiality strategy at time $t-1$, by the combination of monotonicity (as $z_{1j} \succeq r(Z'_{1j})$) and weak-herding (due to (5.12)) resulting in $\sigma_{1j}(z_{1j}, \Omega(z, A, t-1)) \geq \sigma_{1j}(r(Z'_{1j}), \Omega(r(Z'), A, t-1)) = 1$.

It remains to prove that for each $i \neq 1$, Z_{ij} is in the contraction set of Z'_{ij}, or, equivalently, that the depth of $dig(Z_{ij})$ is not greater than the depth of $D'_{ij} = dig(Z'_{ij})$. Note that for weak herding, there is nothing to prove: Changing 1's histories has no effect on the partiality strategies for other products. So the following is for competitive weak herding, and $m = i = 2$:

Assume to the contrary that $dig(Z_{ij})$ is deeper than $dig(Z'_{ij})$. Consider the lemma applied to $r(Z) \in \mathcal{G}_{t-1}$ as above, and note that for $i \neq 1$, $z_{ij} = r(Z_{ij}) \in \mathcal{H}_{t-1}$. By the induction hypothesis, $dig(z_{ij})$ is *not* deeper than $d'_{ij} = dig(r(Z'_{ij}))$. This is possible only if $dig(z_{ij}) = d'_{ij}$, $Z_{ij}(t) \neq N$ and $Z'_{ij}(t) = N$. This entails $\sigma(z_{ij}, \Omega(Z, A, t-1)) = 1$ and $\sigma(r(Z'_{ij}), \Omega(Z', A, t-1)) = 0$. Noting that $r(Z'_{ij}) \succeq z_{ij}$

(because $dig(r(Z'_{ij})) = d'_{ij} = dig(z_{ij}))$ and (5.13), this violates the combination of monotonicity and competitive weak herding[5]: a contradiction.

This proves part 1a and part 1b of the lemma.

The last part (5.11) is a consequence of the first: For each product $i \neq 1$, Z'_{ij} is an implementation of D_{ij}, while Z_{ij} is in the contraction set of D_{ij}. Therefore $Q_i(Z) \geq Q_i(Z')$. Multiplying for each $i \neq 1$ we get $\frac{Q(Z)}{Q_1(Z)} \geq \frac{Q(Z')}{Q_1(Z')}$. □

Continuing to prove our theorem, fix an initial market share vector A and a round t. Define $\mathcal{X}_t(A) \subset \mathcal{G}_t$ to be the set of t-deep history ensembles which are summed in (5.10), i.e.:

$$\mathcal{X}_t(A) := \{Z \in \mathcal{G}_t, \Omega_1(Z, A, t) > \Omega_2(Z, A, t), c(Z, A) = 1\} \tag{5.14}$$

We call two partial history ensembles $Z, Z' \in \mathcal{G}_{\leq t}$ *intersecting*, denoted $Z \backsim Z'$ if the following condition holds: In each pair of corresponding history elements one is in the contraction set of the other i.e. $\forall i \in [m], j \in [n], Z_{ij} \in \Phi^-(Z'_{ij})$ or $Z'_{ij} \in \Phi^-(Z_{ij})$. (Note that $\backsim$ is not an equivalence relation, since it is symmetric, but not transitive).

We now claim that $\mathcal{X}_t(A)$ contains no intersecting pairs $Z \backsim Z'$. Otherwise, there are two corresponding histories in Z and Z' where one has an N-round where the other has either an S- or an F-round. Let us consider, then, the earliest difference between the two history ensembles. Since the partiality strategy is pure, one of these would have zero probability at the point of difference, leading to a zero ex-ante probability for the entire history ensemble. As all $Z \in \mathcal{X}_t(A)$ satisfy $c(Z, A) = 1$ by assumption, this is impossible.

A *spot improvement* of $\mathcal{X}_t(A)$ is a pair of history ensembles (Z', Z) where $Z' \in \mathcal{X}_t(A)$, and $Z \in \mathcal{G}_{\leq t}$ is a partial history ensemble whose digest can be derived from Z''s by (i) changing exactly one F-event of one of $dig(Z')$'s product 1's histories to an S-event, leaving all other product 1 histories unchanged (ii) Optionally contracting all histories other than product 1's, i.e. $\forall i \neq 1, j \in [n], Z_{ij} \in \Phi^-(Z'_{ij})$. By Lemma 4, there exists such a Z with $c(Z, A) = 1$. Due to the derivation of the spot improvement we have $S_1(Z) = S_1(Z') + 1$ and $F_1(Z) + 1 = F_1(Z')$, so, multiplying both sides of (5.11) by $q_1^{S_1(Z')}(1 - q_1)^{F_1(Z')}$, a positive quantity as $0 \leq q_1 \leq 1$, we have:

$$\frac{Q(Z)}{q_1} \geq \frac{Q(Z')}{1 - q_1} \tag{5.15}$$

Let us rephrase this in terms of exclusively non-partial t-deep history ensembles. Z' is already t-deep and not partial. For Z, recall from (5.4) that:

$$Q(Z) = \sum_{Y \in \Psi_t(Z)} Q(Y) \tag{5.16}$$

Therefore:

$$\frac{1}{q_1} \sum_{Y \in \Psi_t(Z)} Q(Y) \geq \frac{1}{1 - q_1} Q(Z') \tag{5.17}$$

All histories in (5.17) are in $\mathcal{X}_t(A)$: Z' is there by definition of the spot improvement. By (5.12) $\forall Y \in \Psi_t(Z), \Omega_1(Y, A, t) \geq \Omega_1(Z', A, t)$, while for all products

[5]It is here that the case $m > 2$ fails. The presence of a 3rd product, whose market share has possibly decreased, would prevent citing competitive weak herding.

other than 1, including product 2, by (5.13) $\Omega_2(Y,A,t) \leq \Omega_2(Z',A,t)$. Therefore $\Omega_1(Y,A,t) > \Omega_2(Y,A,t) \Rightarrow Y \in \mathcal{X}_t(A)$.

Let us sum (5.17) over all possible spot improvements.

As noted above, on both sides this grand total will be over members of $\mathcal{X}_t(A)$. We investigate the multiplicities of each such member on each side:

On the right-hand side, each member $Z \in \mathcal{X}_t(A)$ is summed exactly $F_1(Z)$ times, the number of different ways a 1's history F-event may be changed into an S-event.

On the left-hand side, we claim that each member $Z \in \mathcal{X}_t(A)$ is summed *at most* $S_1(Z)$ times, arguing as following: For Z to appear on the left-hand side, it must have been an extension of (i.e. in the $\Psi_t(\cdot)$ set) of a spot improvement of some other $Z' \in \mathcal{X}_t(A)$. This means it had exactly one of its S-events changed from an F-event in Z'. Now it is impossible for Z to be derived by a change of *the same S-event* from two different $Z', Z'' \in \mathcal{X}_t(A)$, as this would entail $Z' \backsim Z''$: For observe that $Z' \not\backsim Z'' \Rightarrow \Psi_t(Z') \cap \Psi_t(Z'') = \emptyset$, and furthermore no member of $\Psi_t(Z')$ is in the implementation set of any member of $\Psi_t(Z'')$, and vice versa. But $Z' \backsim Z''$ contradicts the observation that no two members of $\mathcal{X}_t(A)$ are intersecting.

Therefore this grand total sum of (5.17) may be written as:

$$\sum_{Z \in \mathcal{X}_t(A)} \frac{S_1(Z)}{q} Q(Z) \geq \sum_{Z \in \mathcal{X}_t(A)} \frac{F_1(Z)}{1-q} Q(Z) \tag{5.18}$$

Which, reminding ourselves of the definition of $\mathcal{X}_t(A)$ (5.14), results in (5.10), which is our theorem restricted to pure partiality strategies. As already shown, this is sufficient to prove (5.6) and the theorem. □

We now state and prove the main result:

Theorem 5. (1) *Let there be m products, and let products $1, 2$ be anonymous but have possibly different qualities q_1, q_2 respectively, with symmetric prior on their quality and initial market share. Let all customers have monotone partiality strategies to these products, and let all customers be weakly herding. Let ω_1, ω_2 be the observed market share after time t of 1 and 2, respectively. Then:*

$$\mathbf{P}[q_1 \geq q_2 | \omega_1 > \omega_2] \geq \mathbf{P}[q_2 \geq q_1 | \omega_1 > \omega_2] \tag{5.19}$$

(2) *The same holds when all customers are competitively weakly herding, rather than weakly herding, and there are two products ($m = 2$).*

Proof. The observed market share is the result of some history ensemble $Z \in \mathcal{G}_t$ and some initial market share vector $A = (A_1, \dots, A_m)$, such that:

$$\omega_1 = \Omega_1(Z,A,t) \tag{5.20}$$

$$\omega_2 = \Omega_2(Z,A,t) \tag{5.21}$$

As the two products are anonymous, and the market share prior is symmetric, we must have $\mathbf{P}[\omega_1 > \omega_2 | q_1 = q_2] = \mathbf{P}[\omega_2 > \omega_1 | q_1 = q_2]$. By theorem 4:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_1 > \omega_2 | q_1 = q_2] \tag{5.22}$$

$$\mathbf{P}[\omega_2 > \omega_1 | q_1 \geq q_2] \leq \mathbf{P}[\omega_2 > \omega_1 | q_1 = q_2] \tag{5.23}$$

Therefore:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_2 > \omega_1 | q_1 \geq q_2] \tag{5.24}$$

By anonymity, $\mathbf{P}[\omega_2 > \omega_1 | q_1 \geq q_2] = \mathbf{P}[\omega_1 > \omega_2 | q_2 \geq q_1]$, therefore:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_1 > \omega_2 | q_2 \geq q_1] \tag{5.25}$$

¿From which the theorem follows by Bayes' theorem and the symmetric prior on q_1, q_2. □

6. Justifying the Assumption of Pure Partiality Strategies

In this section we state and prove a result that is used in previous sections to reduce the proof of certain inequalities to the case of pure partiality strategies. We recall that a *multinear polynomial* is a multivariate polynomial where every monomial is a product of distinct variables.

Theorem 6. *Let $P \subseteq \mathbb{R}^M$ be a polytope that is defined as follows*

$$\forall i \in [M] \qquad 0 \leq x_i \leq 1 \tag{6.1}$$

$$\forall (i,j) \in H \qquad x_i \leq x_j \tag{6.2}$$

for some $H \subseteq [M] \times [M]$. Let G be a graph on $[M]$ where ij is an edge iff $(i,j) \in H$ or $(j,i) \in H$. Let F be a multilinear function on $\mathbb{R}^M$ such that x_i and x_j appear in the same monomial in F only if vertices i and j belong to different connected components of G.

Then $F(x)$ attains its minimum over P at a point x^* all whose coordinates are integers, i.e. $\forall i \in [M], x^*_i \in \{0,1\}$.

Proof. Associated with every $x' := \{x'_1, \ldots, x'_M\} \in P$ and a set $S \subseteq [M]$ is the following subset of P that we call the *segment* of x' and S in P. It is defined by the following equations:

$$\forall i,j \in S \qquad x_i = x_j \tag{6.3}$$

$$\forall i \notin S \qquad x_i = x'_i \tag{6.4}$$

This is easily seen to be the intersection of P and a line and is, therefore, a (possibly empty) one-dimensional sub-polytope of P. It is defined by the requirements that all coordinates in S are equal to each other, while the other coordinates are held constant. We call the value that is common to all coordinates in S the *segment variable*. We say that $F(x)$ is linear over a segment if it is linear in that segment's variable.

Note that a multilinear polynomial need not be linear over a given segment. For example, take the multilinear function $f(x_1, x_2) = x_1 x_2 - x_1$ and the two-dimensional polytope $0 \leq x_1 \leq x_2 \leq 1$. Consider the segment $x_1 = x_2 = y$ where y denotes the segment variable. Note that f is *not* linear but quadratic over this segment[6]. Nevertheless, due to the restrictions placed on $F(x)$ the following proposition holds:

Proposition 1. *Let $x' \in P$ and let S be contained in some connected component C of G. Then $F(x)$ is linear over the segment in P of x' and S.*

Proof. By our assumption about F and since $S \subset C$, every monomial of F contains at most one variable x_i with $i \in S$. The conclusion follows. □

[6]Indeed, in this case, f attains its minimum at non-integral values $x^*_1 = x^*_2 = \frac{1}{2}$

For $x' \in P$ we consider the set of coordinates i for which $1 > x_i > 0$. We divide the set of such coordinates into disjoint *bundles*, where each bundle S is a maximal set with the following properties:

(1) All x'_i with $i \in S$ have the same value $y_S \notin \{0, 1\}$.
(2) The set S is contained in a connected component of G.

Let $x^* \in P$ be a *minimum* of $F(x)$ over P that has the *least* number of bundles. If there are no bundles, then $x^* \in \{0,1\}^M$, as claimed. Otherwise, let S be a bundle. Namely, $S \subseteq C$ for some component C of G, and $x^*_i = y^*$ for each $i \in S$ for some $y^* \notin \{0, 1\}$.

We make several observations on J, the segment in P of x^* and S:

- It is nonempty, since $x^* \in J$.
- The point x^* is not an endpoint of J. Otherwise, there is an inequality (6.2) that x^* satisfies with equality. Namely, $x_j = x_k = y^*$ where $j \in S$ and $k \notin S$. But this implies that $k \in C$ contrary to the maximality of S.
- By Proposition 1 $F(x)$ is linear over the segment.

A linear function that is defined on an interval takes its maximum only at an endpoint unless the function is constant. Therefore, if y is the segment's variable it must appear with a zero coefficient, and $F(x)$ is constant over the segment. Let $y = y_0 \neq y^*$ be the segment variable's value at an endpoint of the segment. The possible values for y_0 are 0, 1 and x^*_k where $k \notin S$. We modify the point x^* by changing, for all $i \in S$, the value of x_i from y^* to y_0. This either eliminates the bundle S or merges it with the bundle of k. Either way the number of bundles is reduced by one, a contradiction. □

7. Discussion and Conclusion

We proved that market share indicates quality in the context of a model where customers base their strategy on their history with products and on market share data, under fairly weak restrictions on their behavior.

One consequence of the result is its guidance to the behavior of the customers themselves: A new customer, with no previous experience of the products, is advised to put her trust in market share data available. In a market in which customer-product interaction is one-shot, and all customers are equally informed about market share, all rational customers should behave alike.

The framework we used in deriving our results can be naturally generalized. While the restrictions of monotonicity and weak herding were successful for reaching the result, we do not claim that our formulation is the only one possible. Different formulations may be attempted, perhaps introducing other considerations into customer strategy. It should be apparent that generalizing our result would require the proof of results similar to Lemmas 2 and 4. We believe that the rest of our proof would carry through.

For example, we believe that the requirement that all customers obey weak herding is too strict, and that some level of elitistic customer behavior does not, in itself, invalidate the result. Namely, so long as elitism is outweighed (in some sense, to be defined) by herding behavior, inferences from market share to quality remain valid.

If elitism becomes the norm, this may not be true, as is illustrated by the following simple albeit artificial example:

Example 1. *Let there be two products* 1 *and* 2*, with* 1 *the superior product:* $q_1 > q_2$*. Let there be* n *customers, divided into two categories. Customer* 1*'s partiality strategy to each product is* 1 *if she has no prior history with the product or if her last experience with it was good, and* 0 *otherwise. Customer* 1*'s strategy ignores market share. Customers* 2 *to* n*, on the other hand, are pure elitists: They will consume a product unconditionally unless that product is a leader in market share, in which case they will not consume it. The initial market share is* $A = \{0, 0\}$*.*

We analyze this to show that, on the 3rd round, market share does not indicate quality.

On the 1st round all customers consume all products. The market share after the round is $\omega = \{n, n\}$*, so there is no market leader.*

On the 2nd round, customers 2 *to* n *will consume all products. Customer* 1*, however, will do so only if her last round was a success. With probability* $q_1(1-q_2)$ *product* 1 *will lead in market share, while with a smaller probability* $q_2(1-q_1)$ *(since* $q_1 > q_2$*) product* 2 *will lead in market share. In other cases (probability* $q_1q_2 + (1-q_1)(1-q_2)$*), market share parity will continue.*

On the 3rd round, if any product has larger market share, the elitists will stop consuming it, but will continue consuming the other product. The result will be that at the end of the round, market leadership will be reversed. As 1 *is the better product, and assuming* n *large enough, the conclusion is that at the end of round* 3 *higher market share indicates lower quality.*

This example is artificial, *inter alia*, in that the negative result depends on the round number. In the example, market share leadership will oscillate, and even rounds will behave differently from odd rounds. We have not been able to find an economically realistic scenario that invalidates our central thesis.

Another generalization is by the introduction of money, which does not play a role in our current model due to the assumption of an undifferentiated market. This may be readily achieved by factoring price into quality, so that customer satisfaction is tied to perceived "value for money". It is not necessary for all customers to have the same sensitivity to price: The generalization ("market share indicates value for money") clearly holds if all customers, were they fully informed about product qualities, agree that a certain product is preferable to another.

We were able to prove the case with competitive weak herding for two products only. Whether the result in fact holds for 3 products or more needs to be clarified. If the answer turns out to be negative, it will be interesting whether an alternative for competitive weak herding exists in which customers are responsive to market share ranking and for which the main result holds for any number of products.

Our analysis assumed that customers are aware of current market share only. What happens if we allow for awareness of historic market share values? There are two aspects to this: First, in the influence of market share data on customer strategies: We defined weak herding as a restriction on customers' response to *current* market share. If this definition of weak herding remains unchanged, then a weak-herding customer, even if aware of historic market share data, may only use it as a tie-breaker in forming her strategies. It should be apparent that this does not disturb the validity of the results, as stated. Second, and however, the statement "market share indicates quality", interpreted as a statement about the probability of better quality conditional on market share data *not* restricted to current data, may be invalid. This is especially true if customers are aware of trends. E.g. The leader in a market may be seen to be losing ground, while the follower is seen to be gaining ground.

We modeled quality as an unchanging attribute of a product. What happens if quality varies between rounds? The main result becomes moot. However it may be asked whether the intermediate results that show that market share is monotonically non-decreasing in the quality (Theorems 1 and 4) remain true. Differentials in the quality may be replaced by partial differentials in any particular round's quality. The answer to this turns out to be negative in the general case. (For observe that the proof of Theorem 1 is based on the fact that all implementations of a history digest Z share the same factor $q^{S(Z)}(1-q)^{F(Z)}$ in their ex-ante probabilities. In particular, Lemma 2 showing the shortening of implementations for improved digests is of no use when the shorter history may have lower probability than a longer one. Moreover, when this factor varies between implementations, a counterexample to any puported equivalent of Theorem 1 may be constructed.)

The significance of Theorem 4 transcends the scope of our main result, applying to differentiated markets where products are differentiated by brand names, price, etc. Though the differentiation means that inferring higher quality from higher market share is not possible, improving quality will always result in higher probability for market share leadership.

References

[1] Anderson, E., C. Fornell and D. R. Lehmann: Customer Satisfaction, Market Share, and Profitability: Findings from Sweden. *Journal of Marketing*, 58: 53-66, 1994.

[2] Ban, A. and N. Linial: The Dynamics of Reputation Systems. *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pp. 91-100, 2011.

[3] Bikhchandani, S., Hirshleifer, D. and Welch, I: A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, **100**, 992-1026, 1992.

[4] Ellison D. and D. Fudenberg Word-of-Mouth Communication and Social Learning, *Quarterly Journal of Economics*, **110(1)**, 93-126. 1995.

[5] Rust, R. and A. Zahorik: Customer Satisfaction, Customer Retention, and Market Share. *Journal of Retailing*, **69** (Summer), 145-156, 1993.

[6] Smallwood, D.E. and J. Conlisk: Product Quality in Markets Where Consumers are Imperfectly Informed. *The Quarterly Journal of Economics*, Vol. 93, No. 1. (Feb., 1979), pp. 1-23.