

Market Share Indicates Quality

Amir Ban

Center for the Study of Rationality, Hebrew University, Jerusalem, Israel

Nati Linial

School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel

Abstract

Market share and quality, or customer satisfaction, go together. Yet inferring one from the other appears difficult. Indeed, such an inference would need detailed information about customer behavior, and might be clouded by modes of behavior such as herding (following popularity) or elitism, where customers avoid popular products. We investigate a fixed-price model where customers are informed about their history with products and about 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 unrestrictive assumptions on customer behavior.

JEL codes: D11, D40, L10.

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 rationality 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.

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 eco-

Email addresses: amirban@netvision.net.il (Amir Ban), nati@cs.huji.ac.il (Nati Linial)

conomic benefit. While not conclusive, empirical studies (e.g. Anderson et. al. [Anderson,1994], Rust and Zahorik [Rust,1993]) generally support a positive correlation between quality and market share.

Market share affects real or perceived quality through various mechanisms, including the creation of network externalities. Yet the direction of this potential impact is ambiguous in the literature. An empirical study by Hellofs and Jacobson[Hellofs,1999] concludes “The view that ...market share and product quality are reinforcing mechanisms ...seems premature”. Our study makes room for both positive and negative impact, through the modeling of “herders” and “elitists”.

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 customers realistically face. In the exceptions, e.g. when offered a Ferrari and a Fiat, price or other differentiation will typically also exist. The extent to which customers may remain unaware of quality has been studied by, e.g., Mitra and Golder[Mitra,2006], who showed that only a fraction of actual quality change registers with customers, and then typically with a delay of years.

Caminal and Vives[Caminal,1996] studied the effect of firms’ signaling of quality on imperfectly informed customers. The signaling is facilitated by strategic price reductions, which increase perceptions of quality by increasing market share. As customers best-respond with their own strategy, in one possible equilibrium market share indicates quality. Our motivation is different. We seek to illuminate the inter-relationship of market share and quality under the least possible restrictions on customer behavior, and independently of signaling, for several reasons: (i) A large body of economic research indicates that customers’ decisions use *bounded* rationality. (ii) Strategic signaling of quality is absent when there is no interest or possibility of doing so, e.g. when prices are set by a regulator (health insurance), or by a non-profit (museum entry tickets), or by a channel (movie tickets), or due to the shortness of the interaction (shore restaurants serving a boatload of tourists stopping for lunch will inevitably have same prices). (iii) The sparseness of theoretical results, and the inconclusiveness of empirical results, increases the value of a basic result. Especially so since, as we will point out, and as the extensive literature on herding shows, there is nothing simple or inevitable about it.

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.

The current work also sets the ground for future work that will include variable prices. A close inspection of our methodology reveals its suitability for such a task. It is particularly encouraging to see its success in reaching general results from a few weak assumptions. One of our results, that the probability for market leadership increases with quality (Theorem 4) applies to variable-price markets. In the Discussion we suggest a path to extending the main result for variable prices.

Though our research is not framed in the canonical terms of action payoffs and Bayesian updating/learning, it in fact conforms to it: A preference for satisfaction over dissatisfaction is all we need to assume about payoffs, and, the monotonicity property, which is introduced in the next section, has a property common to all Bayesian updating rules: A strictly superior sequence of events results in a superior posterior.

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 strategy¹ 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 to 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.

1.2. 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[Smallwood,1979] 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 [Ban,2011] 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[Ellison,1995] 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.[Bikhchandani,1992], consider the inferences that observers can make on the quality of a service based on the customers queuing for that service and how informed those customers are known to be. They show this leads to *herding*, the phenomenon where customers

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

accumulate due to the presence of others. Smallwood and Conlisk[Smallwood,1979] as well as Ban and Linial[Ban,2011] also show that lower quality products can maintain higher market share indefinitely. Lest the reader suspect that the main result is self-evident when customers are Bayesian, since “better quality always wins with Bayesian learning”, note that this assumption is refuted by the above cited papers which are framed in strictly Bayesian terms. This false intuition is dispelled as well by the k -armed bandit problem, where an unlucky start may cause a Bayesian player to miss his best option for a long time or even 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.

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. Conclusions are given in section 6.

Acknowledgement

The original version of this paper was based on a more complicated argument. The proof method that we finally adopted was generously offered by an anonymous referee to whom we are very grateful.

2. Basic Model, When Only History Matters

In our model, customers make decisions regarding products in *rounds* of discrete time $t = 1, 2, \dots$. 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 $\mathcal{H}_t := \{S, N, F\}^t$. The set of all histories are denoted $\mathcal{H} := \bigcup_{k=0}^{\infty} \mathcal{H}_k$. For $Z \in \mathcal{H}_t$ we mark Z 's *depth* $|Z| := t$. Z is composed of *events*, $Z = (Z(1), \dots, Z(t))$ with $Z(k)$ standing for the event in round k . Z^k stands for the sequence of events until round k , i.e. $Z^k = (Z_1, \dots, Z_k) \in \mathcal{H}_k$, so that $Z = Z^t$.

The customer's *partiality strategy*, $\sigma : \mathcal{H} \rightarrow [0, 1]$, 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 $Z \in \mathcal{H}$.

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

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.

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

$$\sigma(Z_1) \geq \sigma(Z_2) \text{ whenever } Z_1 \succeq Z_2.^2 \quad (2.1)$$

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*

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

where the probability space is \mathcal{H} .

PROOF. 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

$$\begin{aligned} Z &\rightarrow ZS && \text{with probability } q \cdot \sigma(Z) \\ Z &\rightarrow ZN && \text{with probability } 1 - \sigma(Z) \\ Z &\rightarrow ZF && \text{with probability } (1 - q) \cdot \sigma(Z) \end{aligned} \quad (3.2)$$

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

$$\mathbf{P}[Z; q] := c(Z)Q(Z; q)$$

²In more general models where the customer's strategy is dependent on more information, e.g. history of other products, or market shares, the generalization of this will be "all else being equal". E.g. $\sigma(Z_1, X, Y, \dots) \geq \sigma(Z_2, X, Y, \dots)$ whenever $Z_1 \succeq Z_2$, where X, Y, \dots denote any other information-set variables. See e.g. the definition of monotonicity in Section 4.

$$Q(Z; q) := q^{S(Z)}(1 - q)^{F(Z)}.$$

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:

$$c(Z^k) = \begin{cases} 1 & k = 0 \\ \sigma(Z^{k-1})c(Z^{k-1}) & Z(k) \neq N \\ [1 - \sigma(Z^{k-1})]c(Z^{k-1}) & Z(k) = N \end{cases} \quad (3.3)$$

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 .

Let $q', q \in [0, 1]$ s.t. $q' > q$. For the proof, we will construct a joint distribution of two types of history:

1. The distribution of histories $\in \mathcal{H}_t$ under product quality q . Histories from this distribution are denoted by Z .
2. The distribution of histories $\in \mathcal{H}_t$ under product quality q' . Histories from this distribution are denoted by Z' .

The joint distribution is defined for the purpose of our analysis. It should be considered a proof technique and not a representation of any real customer behavior.

The construction assigns to each pair of histories $Z, Z' \in \mathcal{H}_t$ a joint probability $f(Z, Z') \in [0, 1]$. Assuming $f(Z, Z')$ to be tabulated in a table whose rows correspond to values of Z and columns correspond to values of Z' , a valid joint distribution must have rows summing to the correct history event probabilities in both rows and columns, i.e. for every $Z \in \mathcal{H}_t$:

$$\sum_{Z' \in \mathcal{H}_t} f(Z, Z') = \mathbf{P}[Z; q] = c(Z)Q(Z; q) \quad (3.4)$$

and for every $Z' \in \mathcal{H}_t$:

$$\sum_{Z \in \mathcal{H}_t} f(Z, Z') = \mathbf{P}[Z'; q'] = c(Z')Q(Z'; q') \quad (3.5)$$

Our construction restricts the joint distribution's support (i.e. those pairs Z, Z' where $f(Z, Z') > 0$) by the following criterion:

Criterion 1. For every $0 < k \leq t$, there exists $0 < l \leq k$ s.t. $\text{dig}(Z'^l) \succeq \text{dig}(Z^k)$.

Consequently, for every $0 < k \leq t$, $\text{con}(Z'^k) \geq \text{con}(Z^k)$, and $\text{con}(Z'^k) = \text{con}(Z^k) \Rightarrow Z'^k \succeq Z^k \Rightarrow \sigma(Z'^k) \geq \sigma(Z^k)$.

We specify the joint distribution in two different ways: In the first, it is specified how the row marginal probabilities, i.e. probabilities of Z events, are split into the row's individual probabilities. The construction will explicitly guarantee (3.4) and Criterion 1. In the second, it is specified how the column marginal probabilities,

i.e. probabilities of Z' events, are split into the column's individual probabilities. The construction will explicitly guarantee (3.5) and Criterion 1. Finally, we will demonstrate that the two constructions coincide and yield the same probabilities, thus proving that it is in fact a joint probability fulfilling (3.4), (3.5) and Criterion 1.

The first, row construction, is specified by:

Let $Z, Z', X, Y \in \mathcal{H}_t$ (we use X in place of Z , and Y in place of Z' for better readability), and $0 < k \leq t$. We define below probability functions $g_1(X, Y, k)$, $h_1(X, Y, k)$. Based on these, the probabilities $f(Z, Z')$ will be given by:

$$g(X, Y, k) := \begin{bmatrix} g_1(X, Y, k) & Y(k) \neq N \\ 1 - g_1(X, Y, k) & Y(k) = N \end{bmatrix} \quad (3.6)$$

$$h(X, Y, k) = \begin{bmatrix} h_1(X, Y, k) & Y(k) = S \\ 1 - h_1(X, Y, k) & Y(k) = F \\ 1 & Y(k) = N \end{bmatrix} \quad (3.7)$$

$$r(X, Y, k) := g(X, Y, k)h(X, Y, k) \quad (3.8)$$

$$R(Z, Z') := \prod_{k=1}^t r(Z, Z', k) \quad (3.9)$$

$$f(Z, Z') := R(Z, Z') \mathbf{P}[Z; q] \quad (3.10)$$

From Equations (3.6) to (3.8) we derive:

$$r(X, Y, k) = \begin{bmatrix} g(X, Y, k)h(X, Y, k) & Y(k) = S \\ g(X, Y, k)[1 - h(X, Y, k)] & Y(k) = F \\ 1 - g(X, Y, k) & Y(k) = N \end{bmatrix} \quad (3.11)$$

The three values of which sum to 1. It follows that for every given values of Z and Z'^{k-1} there holds $\sum r(Z, Z', k) = 1$ where the sum is over $Z'(k) = N, S$ and F . By (3.9) it follows that:

$$\sum_{Z' \in \mathcal{H}_t} R(Z, Z') = 1 \quad (3.12)$$

from which (3.4) follows. Note that this holds regardless of the choice of g_1, h_1 , which is made so as to guarantee that Criterion 1 holds. The definition of g_1 and h_1 follows:

Mark $c := \text{con}(X^{k-1})$, $c' := \text{con}(Y^{k-1})$:

$$g_1(X, Y, k) = \begin{bmatrix} \begin{bmatrix} 1 & X(k) \neq N \\ \frac{\sigma(Y^{k-1}) - \sigma(X^{k-1})}{1 - \sigma(X^{k-1})} & X(k) = N \end{bmatrix} & c = c' \\ \sigma(Y^{k-1}) & c \neq c' \end{bmatrix} \quad (3.13)$$

$$h_1(X, Y, k) = \begin{bmatrix} \begin{bmatrix} 1 & [dig(X)](c+1) = S \\ \frac{q'-q}{1-q} & [dig(X)](c+1) = F \end{bmatrix} & c < con(X) \\ q' & c \geq con(X) \end{bmatrix} \quad (3.14)$$

An examination of (3.6) to (3.14) leads to the following observations:

- The definition of g and g_1 insures that in r 's support $con(Z'^k) \geq con(Z^k)$ for every $k \in [1, t]$: In case Z and Z' have the same consumption at round $k-1$ and $X(k) \neq N$, there is zero probability that $Y(k) = N$. The definition of h and h_1 insures that in r 's support nowhere do $l \in [1, con(Z)]$ $[dig(Z)](l) = S$ and $[dig(Z')](l) = F$ hold simultaneously: If $[dig(X)](c+1) = S$, then with probability zero does $[dig(Y)](c+1) = Y(k)$ equal F . This is exactly our notion of history superiority, denoted by the relation \succeq . Consequently Criterion 1 holds in the support of f .
- As $q' > q$, $0 \leq h(Z, Z', k) \leq 1$, and as Criterion 1 specifies $Z'^k \succeq Z^k$ whenever $con(Z'^k) = con(Z^k)$, $0 \leq g(Z, Z', k) \leq 1$. Therefore $0 \leq R(Z, Z') \leq 1$ for every Z, Z' that satisfy Criterion 1.

In summary, (3.10) defines a matrix $f(Z, Z')$ of non-negative values, which may be non-zero only where Criterion 1 holds, each row summing to the row marginal probability of Z .

We now define a second construction based on columns (values of Z'). The overall structure of this construction is similar to the previous one.

Let $Z, Z', X, Y \in \mathcal{H}_t$, and $0 < k \leq t$. We define below probability functions $g_2(X, Y, k), h_2(X, Y, k)$. Based on these, the probabilities $f'(Z, Z')$ will be given by:

$$g'(X, Y, k) := \begin{bmatrix} g_2(X, Y, k) & X(k) \neq N \\ 1 - g_2(X, Y, k) & X(k) = N \end{bmatrix} \quad (3.15)$$

$$h'(X, Y, k) = \begin{bmatrix} h_2(X, Y, k) & X(k) = S \\ 1 - h_2(X, Y, k) & X(k) = F \\ 1 & X(k) = N \end{bmatrix} \quad (3.16)$$

$$r'(X, Y, k) := g'(X, Y, k)h'(X, Y, k) \quad (3.17)$$

$$R'(Z, Z') := \prod_{k=1}^t r'(Z, Z', k) \quad (3.18)$$

$$f'(Z, Z') := R'(Z, Z') \mathbf{P}[Z'; q'] \quad (3.19)$$

As before, Equations (3.15) to (3.22) yield $\sum r'(Z, Z', k) = 1$ where the sum is over $Z(k) = N, S$ and F . By (3.18) it follows that:

$$\sum_{Z \in \mathcal{H}_t} R'(Z, Z') = 1 \quad (3.20)$$

from which (3.5) follows, irrespective of the choice of g_2 or h_2 , which we define next:

Mark $c := \text{con}(X^{k-1})$, $c' := \text{con}(Y^{k-1})$:

$$g_2(X, Y, k) = \begin{bmatrix} \begin{bmatrix} 0 & Y(k) = N \\ \frac{\sigma(X^{k-1})}{\sigma(Y^{k-1})} & Y(k) \neq N \end{bmatrix} & c = c' \\ \sigma(X^{k-1}) & c \neq c' \end{bmatrix} \quad (3.21)$$

$$h_2(X, Y, k) = \begin{bmatrix} \begin{bmatrix} 0 & [\text{dig}(Y)](c+1) = F \\ \frac{q}{q'} & [\text{dig}(Y)](c+1) = S \end{bmatrix} & c < \text{con}(X) \\ 1 & c \geq \text{con}(X) \end{bmatrix} \quad (3.22)$$

The choice of g_2 guarantees that in r' 's support the consumption requirement of Criterion 1 holds, while the choice of h_2 guarantees that the monotonicity requirement of Criterion 1 also holds. As $q' > q$, $0 \leq h'(Z, Z', k) \leq 1$, and from monotonicity it follows that $0 \leq g'(Z, Z', k) \leq 1$. Therefore $0 \leq R'(Z, Z') \leq 1$ for every Z, Z' that satisfy Criterion 1.

In summary, (3.19) defines a matrix $f'(Z, Z')$ of non-negative values, which may be non-zero only where Criterion 1 holds, each row summing to the column marginal probability of Z .

We now show that both constructions generate the same probabilities, $f(Z, Z') = f'(Z, Z')$, and that therefore the construction constitutes a joint probability distribution whose support satisfies Criterion 1:

We use (3.6) and (3.15) to calculate $\frac{g(Z, Z', k)}{g'(Z, Z', k)}$. Where Criterion 1 holds:³

$$\frac{g(Z, Z', k)}{g'(Z, Z', k)} = \begin{bmatrix} \frac{1-\sigma(Z'^{k-1})}{1-\sigma(Z^{k-1})} & Z'(k) = N, Z(k) = N \\ \frac{1-\sigma(Z'^{k-1})}{\sigma(Z'^{k-1})} & Z'(k) = N, Z(k) \neq N \\ \frac{\sigma(Z'^{k-1})}{1-\sigma(Z^{k-1})} & Z'(k) \neq N, Z(k) = N \\ \frac{\sigma(Z'^{k-1})}{\sigma(Z^{k-1})} & Z'(k) \neq N, Z(k) \neq N \end{bmatrix} = \frac{c(Z'^k)}{c(Z'^{k-1})} \bigg/ \frac{c(Z^k)}{c(Z^{k-1})} \quad (3.23)$$

We define a pairing $[1, t] \mapsto [1, t]$, based on Z, Z' as follows:

- The rounds where Z has a consumption event ($Z(k) \neq N$) are paired with the corresponding rounds of the first $\text{con}(Z)$ consumption events in Z' . I.e. if $\text{con}(Z^{k-1}) = \text{con}(Z'^{k'-1}) = l-1$, $\text{con}(Z^k) = \text{con}(Z'^{k'}) = l$, then k is paired with k' .
- Any $\text{con}(Z') - \text{con}(Z)$ N -rounds of Z are paired with rounds of the $[\text{con}(Z) + 1]$ 'th to $\text{con}(Z')$ 'th consumption event in Z' .
- The remaining $t - \text{con}(Z')$ rounds of Z and Z' are all N -rounds, and are paired in any order.

³In the case $c = c'$, $Z(k) \neq N$, $Z'(k) = N$ this expression evaluates to $\frac{0}{0}$, but this case is ruled out by Criterion 1.

We use (3.7) and (3.16) to calculate $\frac{h(Z, Z', k)}{h'(Z, Z', k')}$ where $k \rightarrow k'$ is part of the above pairing. Where Criterion 1 holds:

$$\frac{h(Z, Z', k)}{h'(Z, Z', k')} = \left[\begin{array}{cc} \left[\begin{array}{c} 1 \quad Z(k) = N \\ \frac{1-q'}{1-q} \quad Z(k) = F \\ 1 - q' \quad Z(k) = N \end{array} \right] & \begin{array}{c} Z'(k') = N \\ Z'(k') = F \end{array} \\ \left[\begin{array}{c} \frac{q'}{1-q} \quad Z(k) = F \\ \frac{q'}{q} \quad Z(k) = S \\ q' \quad Z(k) = N \end{array} \right] & \begin{array}{c} Z'(k') = S \end{array} \end{array} \right] = \frac{Q(Z'^{k'}; q')}{Q(Z'^{k'-1}; q')} / \frac{Q(Z^k; q)}{Q(Z^{k-1}; q)} \quad (3.24)$$

Combining (3.9), (3.18), (3.23) and (3.24), we calculate:

$$\frac{R(Z, Z')}{R'(Z, Z')} = \frac{c(Z')}{c(Z)} \frac{Q(Z'; q')}{Q(Z; q)} = \frac{\mathbf{P}[Z'; q']}{\mathbf{P}[Z; q]} \quad (3.25)$$

Recalling (3.10), (3.19) we conclude $f(Z, Z') = f'(Z, Z')$. Therefore the two constructions are the same, as we sought to show, and therefore form a joint distribution with support given by Criterion 1.

The theorem now follows: In the joint distribution matrix, consider a minor with rows for which $\text{con}(Z) \geq x$, and columns for which $\text{con}(Z') \geq x$. Since Criterion 1 requires $\text{con}(Z') \geq \text{con}(Z)$, the minor contains all the support of each included row. Therefore:

$$\sum_{\text{con}(Z) \geq x} \mathbf{P}[Z; q] \leq \sum_{\text{con}(Z') \geq x} \mathbf{P}[Z'; q'] \quad (3.26)$$

Since this is true for any $q' > q$, $\mathbf{P}[\text{con}(Z) \geq x | Z \in \mathcal{H}_t]$ is non-decreasing in the product quality, as the theorem asserts. \square

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 :*

$$\frac{d}{dq} \mathbb{E}[\text{con}(Z) | Z \in \mathcal{H}_t] \geq 0 \quad (3.27)$$

where the probability space is \mathcal{H} .

PROOF. Since:

$$\mathbb{E}[\text{con}(Z) | Z \in \mathcal{H}_t] = \sum_{x=1}^{\infty} \mathbf{P}[\text{con}(Z) \geq x | Z \in \mathcal{H}_t] \quad (3.28)$$

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 (here defined as 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 :*

$$\frac{d}{dq} \mathbf{P} \left[\sum_{j=1}^n \text{con}(Z_j) \geq x | \forall j \in [n], Z_j \in \mathcal{H}_t \right] \geq 0 \quad (3.29)$$

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 $a_n(x) := \mathbf{P} \left[\sum_{j=1}^n \text{con}(Z_j) \geq x | \forall j \in [n], Z_j \in \mathcal{H}_t \right]$ and $b(y) := \mathbf{P} [Z_n \geq y | Z_n \in \mathcal{H}_t]$. Then:

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

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

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

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:*

$$\mathbf{P}\left[q_1 \geq q_2 \mid \sum_{j=1}^n \text{con}(Z_{1j}) > \sum_{j=1}^n \text{con}(Z_{2j})\right] \geq \mathbf{P}\left[q_2 \geq q_1 \mid \sum_{j=1}^n \text{con}(Z_{1j}) > \sum_{j=1}^n \text{con}(Z_{2j})\right]$$

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

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

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

$$\mathbf{P}[\omega_1 > \omega_2 \mid q_1 \geq q_2] \geq \mathbf{P}[\omega_1 > \omega_2 \mid q_1 = q_2] \quad (3.30)$$

$$\mathbf{P}[\omega_2 > \omega_1 \mid q_1 \geq q_2] \leq \mathbf{P}[\omega_2 > \omega_1 \mid q_1 = q_2] \quad (3.31)$$

Therefore:

$$\mathbf{P}[\omega_1 > \omega_2 \mid q_1 \geq q_2] \geq \mathbf{P}[\omega_2 > \omega_1 \mid q_1 \geq q_2] \quad (3.32)$$

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

$$\mathbf{P}[\omega_1 > \omega_2 \mid q_1 \geq q_2] \geq \mathbf{P}[\omega_1 > \omega_2 \mid q_2 \geq q_1] \quad (3.33)$$

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

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. q_i denotes product i 's quality. 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 .

Given a t -deep history ensemble Z , and an initial market share $A := (A_1, \dots, A_m)$, the market share of product $i \in [m]$ after round k is the total number of units con-

summed of product m up to round k , and is denoted by $\Omega_i(Z, A, k)$:

$$\Omega_i(Z, A, k) = A_i + \sum_{j=1}^n \text{con}(Z_{ij}^k) \quad (4.1)$$

All customers are aware of the round- k market share of all products when they make their consumption decisions at round $k + 1$.⁴

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

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.

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 .

Mark $q := (q_1, \dots, q_m)$. Define:

$$S_i(Z) := \sum_{j=1}^n S(Z_{ij}) \quad (4.2)$$

$$F_i(Z) := \sum_{j=1}^n F(Z_{ij}) \quad (4.3)$$

$$Q_i(Z; q) := q_i^{S_i(Z)} (1 - q_i)^{F_i(Z)} \quad (4.4)$$

$$Q(Z; q) := \prod_{i=1}^m Q_i(Z; q) \quad (4.5)$$

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.

A customer j is called *weakly herding* if for every product i , time t , t -deep history Z_{ij} and market share vector $\omega = \{\omega_1, \dots, \omega_m\}$, $\sigma_{ij}(Z_{ij}, \omega)$ is non-decreasing

⁴See the Discussion for some comments on how things behave when customers may have more information about past market shares.

in ω_i (and is independent of ω_k for $k \neq i$). A customer j is called *competitively weakly herding* if for every product i , time t , t -deep history Z_{ij} and market share vector $\omega = \{\omega_1, \dots, \omega_m\}$, $\sigma_{ij}(Z_{ij}, \omega)$ 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 competitively 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 $\omega = (\omega_1, \omega_2, \dots, \omega_m)$. Define the permutation $K_{12}(\omega) := (\omega_2, \omega_1, \omega_3, \dots, \omega_m)$. Then products 1, 2 are *anonymous* if, for each customer j , for each history $Z \in \mathcal{H}$ and for each market share vector ω :

$$\sigma_{1j}(Z, \omega) = \sigma_{2j}(Z, K_{12}(\omega)) \quad (4.6)$$

5. Theorem with Market Share Observation

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

$$\frac{d}{dq_1} \mathbf{P} \left[\Omega_1(Z, A, t) > \Omega_2(Z, A, t) \mid Z \in \mathcal{G}_t \right] \geq 0 \quad (5.1)$$

where $q_i \in [0, 1]$ is the quality of product i for $i = 1, \dots, m$, 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, with changes:

Let $q := (q_1, q_2, \dots, q_m)$ and $q' := (q'_1, q'_2, \dots, q'_m)$ be two product quality vectors satisfying $q'_1 > q_1$ and $q_2 = q'_2, \dots, q_m = q'_m$, so that q' is different from q only in having better quality for product 1. For the proof, we will construct a joint distribution of two types of history ensembles:

1. The distribution of history ensembles $\in \mathcal{G}_t$ under product qualities q . Ensembles from this distribution are denoted by Z .

2. The distribution of history ensembles $\in \mathcal{G}_t$ under product qualities q' . Ensembles from this distribution are denoted by Z' .

As before, the joint distribution is a proof technique and not a representation of any real customer behavior.

The construction assigns to each pair of ensembles $Z, Z' \in \mathcal{G}_t$ a joint probability $f(Z, Z') \in [0, 1]$. Assuming $f(Z, Z')$ to be tabulated in a table whose rows correspond to values of Z and columns correspond to values of Z' , a valid joint distribution must have rows summing to the correct ensemble event probabilities in both rows and columns, i.e. for every $Z \in \mathcal{G}_t$:

$$\sum_{Z' \in \mathcal{G}_t} f(Z, Z') = \mathbf{P}[Z; q] = c(Z)Q(Z; q) \quad (5.2)$$

and for every $Z' \in \mathcal{G}_t$:

$$\sum_{Z \in \mathcal{G}_t} f(Z, Z') = \mathbf{P}[Z'; q'] = c(Z')Q(Z'; q') \quad (5.3)$$

Our construction restricts the joint distribution's support (i.e. those pairs Z, Z' where $f(Z, Z') > 0$) by the following criterion:

Criterion 2. *For every customer $j \in [n]$, and round $0 < k \leq t$:*

- *There exists $0 < l \leq k$ s.t. $\text{dig}(Z'_{1j})^l \succeq \text{dig}(Z_{1j}^k)$.*
- *For every product $i \in [m]$, $i \neq 1$, there exists $0 < l \leq k$ s.t. $\text{dig}(Z'_{ij})^l = \text{dig}(Z_{ij}^k)$. (In the case of weak herding, a fortiori, $Z_{ij}^k = Z'_{ij}^k$).*

The following statements are all consequences of the criterion:

- *$\text{con}(Z'_{1j})^k \geq \text{con}(Z_{1j}^k)$, while for every $i \neq 1$, $\text{con}(Z'_{ij})^k \leq \text{con}(Z_{ij}^k)$*
- *$\text{con}(Z'_{1j})^k = \text{con}(Z_{1j}^k) \Rightarrow Z'_{1j}^k \succeq Z_{1j}^k$, while for every $i \neq 1$, $\text{con}(Z'_{ij})^k = \text{con}(Z_{ij}^k) \Rightarrow Z_{ij}^k \succeq Z'_{ij}^k$.*
- *$\Omega_1(Z', A, k) \geq \Omega_1(Z, A, k)$, while for every $i \neq 1$, $\Omega_i(Z', A, k) \leq \Omega_i(Z, A, k)$.*
- *If monotonicity and either weak herding or competitively weak herding hold, $\sigma_{1j}(Z', \Omega(Z', A, k)) \geq \sigma_{1j}(Z, \Omega(Z, A, k))$.*
- *If monotonicity and either weak herding (for any number of products) or competitively weak herding (for two products) hold, $\sigma_{ij}(Z', \Omega(Z', A, k)) \leq \sigma_{ij}(Z, \Omega(Z, A, k))$. Note that it is here that the case $m > 2$ of competitively weak herding breaks, since the presence of a 3rd product, whose market share possibly decreased from Z to Z' , would not allow drawing this consequence.*

We specify the joint distribution in two different ways: In the first, it is specified how the row marginal probabilities, i.e. probabilities of Z events, are split into the row's individual probabilities. The construction will explicitly guarantee (5.2) and

Criterion 2. In the second, it is specified how the column marginal probabilities, i.e. probabilities of Z' events, are split into the column's individual probabilities. The construction will explicitly guarantee (5.3) and Criterion 2. Finally, we will demonstrate that the two constructions coincide and lead to the same probabilities, thus proving that it is in fact a joint probability fulfilling (5.2), (5.3) and Criterion 2.

The first, row construction, is specified by:

Let $Z, Z', X, Y \in \mathcal{G}_t$ (to improve readability X will be used in place of Z , and Y in place of Z'), $i \in [m]$, $j \in [n]$ and $0 < k \leq t$. We will define below probability functions $g_3(X, Y, i, j, k)$, $h_3(X, Y, i, j, k)$. Based on these, the probabilities $f(Z, Z')$ will be given by:

$$g(X, Y, i, j, k) := \begin{bmatrix} g_3(X, Y, i, j, k) & Y_{ij}(k) \neq N \\ 1 - g_3(X, Y, i, j, k) & Y_{ij}(k) = N \end{bmatrix} \quad (5.4)$$

$$h(X, Y, i, j, k) = \begin{bmatrix} h_3(X, Y, i, j, k) & Y_{ij}(k) = S \\ 1 - h_3(X, Y, i, j, k) & Y_{ij}(k) = F \\ 1 & Y_{ij}(k) = N \end{bmatrix} \quad (5.5)$$

$$r(X, Y, i, j, k) := g(X, Y, i, j, k)h(X, Y, i, j, k) \quad (5.6)$$

$$R(Z, Z') := \prod_{i=1}^m \prod_{j=1}^n \prod_{k=1}^t r(Z, Z', i, j, k) \quad (5.7)$$

$$f(Z, Z') := R(Z, Z') \mathbf{P}[Z; q] \quad (5.8)$$

From Equations (5.4) to (5.6) we derive:

$$r(X, Y, i, j, k) = \begin{bmatrix} g(X, Y, i, j, k)h(X, Y, i, j, k) & Y_{ij}(k) = S \\ g(X, Y, i, j, k)[1 - h(X, Y, i, j, k)] & Y_{ij}(k) = F \\ 1 - g(X, Y, i, j, k) & Y_{ij}(k) = N \end{bmatrix} \quad (5.9)$$

The three values of which sum to 1. It follows that for every given values of Z_{ij} and $Z'_{ij}{}^{k-1}$ there holds $\sum r(Z, Z', i, j, k) = 1$ where the sum is over $Z'_{ij} = N, S$ and F . By (5.7) it follows that:

$$\sum_{Z' \in \mathcal{G}_t} R(Z, Z') = 1 \quad (5.10)$$

from which (5.2) follows. Note that this holds regardless of the specific choice of g_3, h_3 , which come in next.

Mark $c := \text{con}(X_{ij}^{k-1})$, $c' := \text{con}(Y_{ij}^{k-1})$, $\Sigma(Z, k) := \sigma_{ij}(Z_{ij}^k, \Omega(Z, A, k))$

$$g_3(X, Y, i, j, k) = \begin{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & X_{ij}(k) \neq N \\ \frac{\Sigma(Y, k-1) - \Sigma(X, k-1)}{1 - \Sigma(X, k-1)} & X_{ij}(k) = N \end{bmatrix} & i = 1 \\ 0 & X_{ij}(k) = N \\ \frac{\Sigma(Y, k-1)}{\Sigma(X, k-1)} & X_{ij}(k) \neq N \end{bmatrix} & i \neq 1 \end{bmatrix} \begin{matrix} c = c' \\ c \neq c' \end{matrix} \quad (5.11)$$

$$h_{31}(X, Y, i, j, k) = \left[\begin{array}{c} \left[\begin{array}{cc} 1 & [dig(X_{ij}))(c+1) = S \\ \frac{q'_i - q_i}{1 - q_i} & [dig(X_{ij}))(c+1) = F \end{array} \right] \\ q'_i \end{array} \right] \begin{array}{l} c < con(X_{ij}) \\ c \geq con(X_{ij}) \end{array} \quad (5.12)$$

$$h_{3i}(X, Y, i, j, k) = \left[\begin{array}{c} \left[\begin{array}{cc} 1 & [dig(X_{ij}))(c+1) = S \\ \frac{q'_i - q_i}{1 - q_i} & [dig(X_{ij}))(c+1) = F \end{array} \right] \\ 1 \end{array} \right] \begin{array}{l} c < con(Y_{ij}) \\ c \geq con(Y_{ij}) \end{array} \quad (5.13)$$

$$h_3(X, Y, i, j, k) = \left[\begin{array}{cc} h_{31}(X, Y, i, j, k) & i = 1 \\ h_{3i}(X, Y, i, j, k) & i \neq 1 \end{array} \right] \quad (5.14)$$

An examination of (5.4) to (5.14) leads to the following observations:

- The definition of g and g_3 insures that in r 's support $con(Z'_{1j}^k) \geq con(Z_{1j}^k)$ for every $j \in [n]$, $k \in [1, t]$, and $con(Z'_{ij}^k) \leq con(Z_{ij}^k)$ for every $i \in [2, m]$, $j \in [n]$, $k \in [1, t]$. The definition of h and h_3 insures that in r 's support nowhere do $l \in [1, con(Z)]$ $[dig(Z)](l) = S$ and $[dig(Z')](l) = F$ hold simultaneously. This is the definition of history superiority, denoted by the relation \succeq . Consequently Criterion 2 holds in the support of f .
- The difference between (5.12) and (5.13) (compare with (3.14) and (3.22)) stems from the fact that for product 1, X 's digest is shorter or equal to Y 's, while for other products, it is Y 's digest that is shorter or equal.
- As $q'_i \geq q_i$ for every $i \in [m]$, $0 \leq h(Z, Z', i, j, k) \leq 1$. Also, as a consequence of Criterion 2 is that $\Sigma(Z', k-1) \geq \Sigma(Z, k-1)$ for product 1 ($i = 1$), and $\Sigma(Z', k-1) \leq \Sigma(Z, k-1)$ for other products ($i \neq 1$), $0 \leq g(Z, Z', i, j, k) \leq 1$. Therefore $0 \leq R(Z, Z') \leq 1$ for every Z, Z' satisfying Criterion 2.

In summary, (5.8) defines a matrix $f(Z, Z')$ of non-negative values, which may be non-zero only where Criterion 2 holds, each row summing to the row marginal probability of Z .

We now define a second construction based on columns (values of Z'):

Let $Z, Z', X, Y \in \mathcal{G}_t$, $i \in [m]$, $j \in [n]$ and $0 < k \leq t$. We define below probability functions $g_4(X, Y, i, j, k)$, $h_4(X, Y, i, j, k)$. Based on these, the probabilities

$f(Z, Z')$ will be given by:

$$g'(X, Y, i, j, k) := \begin{bmatrix} g_4(X, Y, i, j, k) & X_{ij}(k) \neq N \\ 1 - g_4(X, Y, i, j, k) & X_{ij}(k) = N \end{bmatrix} \quad (5.15)$$

$$h'(X, Y, i, j, k) = \begin{bmatrix} h_4(X, Y, i, j, k) & X_{ij}(k) = S \\ 1 - h_4(X, Y, i, j, k) & X_{ij}(k) = F \\ 1 & X_{ij}(k) = N \end{bmatrix} \quad (5.16)$$

$$r'(X, Y, i, j, k) := g'(X, Y, i, j, k)h'(X, Y, i, j, k) \quad (5.17)$$

$$R'(Z, Z') := \prod_{i=1}^m \prod_{j=1}^n \prod_{k=1}^t r'(Z, Z', i, j, k) \quad (5.18)$$

$$f'(Z, Z') := R'(Z, Z') \mathbf{P}[Z'; q'] \quad (5.19)$$

An examination of (5.15) to (5.24) leads to observations similar to those noted above for the first construction: There holds $\sum r'(Z, Z', k) = 1$ where the sum is over $Z_{ij}(k) = N, S$ and F . By (5.18) it follows that:

$$\sum_{Z \in \mathcal{G}_t} R'(Z, Z') = 1 \quad (5.20)$$

from which (5.3) follows. We define the functions g_4 and h_4 :

Mark $c := \text{con}(X_{ij}^{k-1})$, $c' := \text{con}(Y_{ij}^{k-1})$, $\Sigma(Z, k) := \sigma_{ij}(Z_{ij}^k, \Omega(Z, A, k))$

$$g_4(X, Y, i, j, k) = \begin{bmatrix} \begin{bmatrix} 0 & Y_{ij}(k) = N \\ \frac{\Sigma(X, k-1)}{\Sigma(Y, k-1)} & Y_{ij}(k) \neq N \end{bmatrix} & i = 1 \\ \begin{bmatrix} 1 & Y_{ij}(k) \neq N \\ \frac{\Sigma(X, k-1) - \Sigma(Y, k-1)}{1 - \Sigma(Y, k-1)} & Y_{ij}(k) = N \end{bmatrix} & i \neq 1 \\ \Sigma(X, k-1) & \end{bmatrix} \begin{matrix} c = c' \\ c \neq c' \end{matrix} \quad (5.21)$$

$$h_{41}(X, Y, i, j, k) = \begin{bmatrix} \begin{bmatrix} 0 & [\text{dig}(Y_{ij})](c+1) = F \\ \frac{q_i}{q'_i} & [\text{dig}(Y_{ij})](c+1) = S \end{bmatrix} & c < \text{con}(X_{ij}) \\ 1 & c \geq \text{con}(X_{ij}) \end{bmatrix} \quad (5.22)$$

$$h_{4i}(X, Y, i, j, k) = \begin{bmatrix} \begin{bmatrix} 0 & [\text{dig}(Y_{ij})](c+1) = F \\ \frac{q_i}{q'_i} & [\text{dig}(Y_{ij})](c+1) = S \end{bmatrix} & c < \text{con}(Y_{ij}) \\ q_i & c \geq \text{con}(Y_{ij}) \end{bmatrix} \quad (5.23)$$

$$h_4(X, Y, i, j, k) = \begin{bmatrix} h_{41}(X, Y, i, j, k) & i = 1 \\ h_{4i}(X, Y, i, j, k) & i \neq 1 \end{bmatrix} \quad (5.24)$$

The choice of g_4 guarantees that in r' 's support the market share requirements of Criterion 2 holds, while the choice of h_4 guarantees that the monotonicity re-

quirements of Criterion 2 also holds. As $q'_i \geq q_i$, $0 \leq h'(Z, Z', i, j, k) \leq 1$, and from monotonicity and (competitive) weak herding it follows that $0 \leq g'(Z, Z', i, j, k) \leq 1$. Therefore $0 \leq R'(Z, Z') \leq 1$ for every Z, Z' satisfying Criterion 2.

In summary, (5.15) through (5.24) define a matrix $f'(Z, Z')$ of non-negative values, which may be non-zero only where Criterion 2 holds, each column summing to the column marginal probability of Z' .

We now show that both constructions generate the same probabilities, $f(Z, Z') = f'(Z, Z')$, and that therefore the construction is for a joint probability with Criterion 2 for its support:

We use (5.4) and (5.15) to calculate that where Z, Z' obey Criterion 2:⁵

$$\frac{g(Z, Z', i, j, k)}{g'(Z, Z', i, j, k)} = \left[\begin{array}{ll} \frac{1-\Sigma(Z', k-1)}{1-\Sigma(Z, k-1)} & Z'_{ij}(k) = N, Z_{ij}(k) = N \\ \frac{1-\Sigma(Z', k-1)}{\Sigma(Z, k-1)} & Z'_{ij}(k) = N, Z_{ij}(k) \neq N \\ \frac{\Sigma(Z', k-1)}{1-\Sigma(Z, k-1)} & Z'_{ij}(k) \neq N, Z_{ij}(k) = N \\ \frac{\Sigma(Z', k-1)}{\Sigma(Z, k-1)} & Z'_{ij}(k) \neq N, Z_{ij}(k) \neq N \end{array} \right] = \frac{c_{ij}(Z'^k_{ij})}{c_{ij}(Z'^{k-1}_{ij})} / \frac{c_{ij}(Z^k_{ij})}{c_{ij}(Z^{k-1}_{ij})} \quad (5.25)$$

For each $i \in [m]$, $j \in [n]$, we define a pairing $[1, t] \mapsto [1, t]$, based on Z_{ij}, Z'_{ij} . The following describes the pairing for product 1:

- The rounds where Z_{ij} has a consumption event ($Z_{ij}(k) \neq N$) are paired with the corresponding rounds of the first $\text{con}(Z_{ij})$ consumption events in Z'_{ij} . I.e. if $\text{con}(Z^{k-1}_{ij}) = \text{con}(Z'^{k'-1}_{ij}) = l-1$, $\text{con}(Z^k_{ij}) = \text{con}(Z'^{k'}_{ij}) = l$, then k is paired with k' .
- Any $\text{con}(Z'_{ij}) - \text{con}(Z_{ij})$ N -rounds of Z_{ij} are paired with rounds of the $[\text{con}(Z_{ij}) + 1]$ 'th to $\text{con}(Z'_{ij})$ 'th consumption event in Z'_{ij} .
- The remaining $t - \text{con}(Z'_{ij})$ rounds of Z_{ij} and Z'_{ij} are all N -rounds, and are paired in any order.

For products other than 1, use the above pairing, but exchange Z_{ij} for Z'_{ij} , and vice versa (the reason being that it is Z'_{ij} instead of Z_{ij} that has the shorter digest).

We use (5.5) and (5.16) to calculate that where Z, Z' obey Criterion 2, and

⁵i.e. except for the cases $c = c', i = 1, Z(k) \neq N, Z'(k) = N$ and $c = c', i \neq 1, Z(k) = N, Z'(k) \neq N$, where the expression evalutes to $\frac{0}{0}$

$k \rightarrow k'$ is part of the above pairing:

$$\frac{h(Z, Z', i, j, k')}{h'(Z, Z', i, j, k)} = \left[\begin{array}{c} \left[\begin{array}{cc} \frac{1}{1-q_i} & Z_{ij}(k) = F \\ \frac{1}{q_i} & Z_{ij}(k) = S \\ 1 & Z_{ij}(k) = N \end{array} \right] & Z'_{ij}(k') = N \\ \left[\begin{array}{cc} \frac{1-q'_i}{1-q_i} & Z_{ij}(k) = F \\ \frac{1-q'_i}{q_i} & Z_{ij}(k) = S \\ 1 - q'_i & Z_{ij}(k) = N \end{array} \right] & Z'_{ij}(k') = F \\ \left[\begin{array}{cc} \frac{q'_i}{1-q_i} & Z_{ij}(k) = F \\ \frac{q'_i}{q_i} & Z_{ij}(k) = S \\ q'_i & Z_{ij}(k) = N \end{array} \right] & Z'_{ij}(k') = S \end{array} \right] = \frac{Q(Z'_{ij}^{k'}; q'_i)}{Q(Z'_{ij}^{k'-1}; q'_i)} / \frac{Q(Z_{ij}^k; q_i)}{Q(Z_{ij}^{k-1}; q_i)} \quad (5.26)$$

Combining (5.7), (5.18), (5.25) and (5.26), we calculate:

$$\frac{R(Z, Z')}{R'(Z, Z')} = \frac{c(Z')}{c(Z)} \frac{Q(Z'; q')}{Q(Z; q)} = \frac{\mathbf{P}[Z'; q']}{\mathbf{P}[Z; q]} \quad (5.27)$$

Recalling (5.8), (5.19) we conclude $f(Z, Z') = f'(Z, Z')$. Therefore the two constructions are the same, as we sought to show, and therefore form a joint distribution with support given by Criterion 2.

The theorem now follows: In the joint distribution matrix, consider a minor with rows for which $\Omega_1(Z, A, t) > \Omega_2(Z, A, t)$, and columns for which $\Omega_1(Z', A, t) > \Omega_2(Z', A, t)$. Since a consequence of Criterion 2 is that $\Omega_1(Z', A, t) \geq \Omega_1(Z, A, t)$ and $\Omega_2(Z', A, t) \leq \Omega_2(Z, A, t)$, the minor contains all the support of each included row. Therefore:

$$\sum_{\Omega_1(Z, A, t) > \Omega_2(Z, A, t)} \mathbf{P}[Z; q] \leq \sum_{\Omega_1(Z', A, t) > \Omega_2(Z', A, t)} \mathbf{P}[Z'; q'] \quad (5.28)$$

Since this is true for any $q'_1 > q_1$, $\mathbf{P}[\Omega_1(Z, A, t) > \Omega_2(Z, A, t) | Z \in \mathcal{G}_t]$ is non-decreasing in q_1 , as the theorem asserts. \square

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:*

$$\mathbf{P}[q_1 \geq q_2 | \omega_1 > \omega_2] \geq \mathbf{P}[q_2 \geq q_1 | \omega_1 > \omega_2] \quad (5.29)$$

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:

$$\omega_1 = \Omega_1(Z, A, t) \quad (5.30)$$

$$\omega_2 = \Omega_2(Z, A, t) \quad (5.31)$$

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:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_1 > \omega_2 | q_1 = q_2] \quad (5.32)$$

$$\mathbf{P}[\omega_2 > \omega_1 | q_1 \geq q_2] \leq \mathbf{P}[\omega_2 > \omega_1 | q_1 = q_2] \quad (5.33)$$

Therefore:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_2 > \omega_1 | q_1 \geq q_2] \quad (5.34)$$

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:

$$\mathbf{P}[\omega_1 > \omega_2 | q_1 \geq q_2] \geq \mathbf{P}[\omega_1 > \omega_2 | q_2 \geq q_1] \quad (5.35)$$

from which the theorem follows by Bayes' theorem and the symmetric prior on q_1, q_2 .

6. 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 definition of criteria similar to Criteria 1 and 2, and constructing distributions whose support adhere to these new criteria.

For example, we believe that the requirement that all customers obey weak herding is too strict, and that some level of elitist 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 competitively 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 competitively 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, e.g., that the pairing used in (3.24) means that product qualities from different periods are compared. This may cause the construction to fail by having negative joint probabilities. Interestingly, this problem does not exist, and so our results still hold, if product qualities are variable but non-increasing in time).

References

- [Anderson,1994] ANDERSON, E., C. FORNELL AND D. R. LEHMANN: Customer Satisfaction, Market Share, and Profitability: Findings from Sweden. *Journal of Marketing*, 58: 53-66, 1994.
- [Ban,2011] 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.
- [Bikhchandani,1992] 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.
- [Caminal,1996] CAMINAL, R., AND VIVES, X.: Why market shares matter: an information-based theory. *The RAND Journal of Economics*, 221-239, 1996.
- [Ellison,1995] ELLISON D., AND D. FUDENBERG Word-of-Mouth Communication and Social Learning, *Quarterly Journal of Economics*, **110(1)**, 93-126. 1995.
- [Hellofs,1999] HELLOFS, L. L., AND JACOBSON, R.: Market share and customers' perceptions of quality: when can firms grow their way to higher versus lower quality? *The Journal of Marketing*, 16-25, 1999.
- [Mitra,2006] MITRA, D., AND GOLDER, P. N. How does objective quality affect perceived quality? Short-term effects, long-term effects, and asymmetries. *Marketing Science*, **25(3)**, 230-247, 2006.
- [Rust,1993] RUST, R., AND A. ZAHORIK: Customer Satisfaction, Customer Retention, and Market Share. *Journal of Retailing*, **69** (Summer), 145-156, 1993.
- [Smallwood,1979] 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.