A Marketing Game: a Model for Social Media Mining and Manipulation

Matthew G. Reyes

Independent Researcher and Consultant,
Ann Arbor MI 48105, USA,
matthewgreyes@yahoo.com,
matthewgreyes.com,
amarketinggame.com

Abstract. This paper derives marketing-influenced Glauber dynamics for socially-contingent consumer choice, which rests on the foundation of socially-contingent random utility. This dynamics model provides companies with a reinforcement learning approach to influencing consumer decision-making. The paper presents a procedure for using machine learning algorithms to estimate consumer preferences as well as direct and social biases on the network. The paper discusses the use of market research to estimate inherent biases and marketing responses for individual consumers. Finally, the paper illustrates on a star-chain network how optimization of marketing allocation depends on parameter estimation.

1 A Marketing Game

This is principally a position paper in which we argue for a model of consumer decision-making that affords analysis and optimization of marketing influence. Our model builds off of well-established models in economics, and while it differs in subtle points, it is precisely these points that open the door for understanding and application of influence on decision-making within a social network.

Consider a *market* in which consumers choose between two alternatives, Product A and Product B, according to their perception of the value of these two choices. The Products may be commercial products or political candidates, for example, an individual's perceived utility of such owing to enhanced productivity, enjoyment, or status. In exchange for the product they choose, they give to Company A or Company B, respectively, their money, or vote, for example. To enhance the perception that consumers have of their respective products, Companies A and B *market* their products to consumers in a social network, as depicted in Figure 1. The point of emphasizing the social network is to underscore the role that social connections play in influencing the decisions of consumers, and the strategy involved in companies selecting which consumers to target with

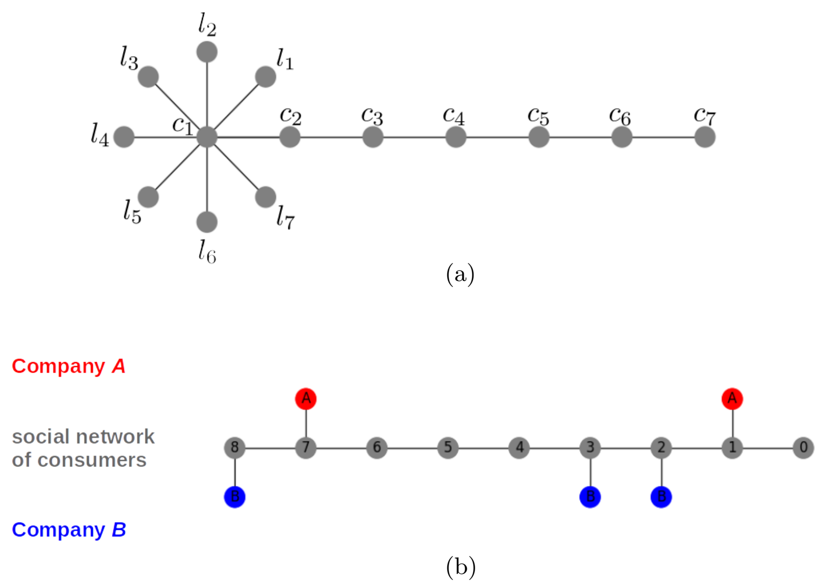

Fig. 1: (a) (perpendicular) view of a star-chain network that does not show marketing. The consumer with $d+1$ neighbors is the *hub* c_1; the d consumers whose only neighbor is the hub are the *leaves* $l_1, \ldots, l_d$; the remaining consumers are the *chain* consumers $c_2, \ldots, c_7$. (b) (parallel) view of network illustrating marketing influence by Company A (red) and Company B (blue).

marketing. One may approach this problem with the *scientific* objective of seeking to characterize influence imparted by particular (types of) individuals or a particular network topology; or the *engineering* objective of seeking to optimize influence irrespective of the respective individuals or topology. This paper addresses the engineering objective.

The market share [3] for Company A with respect to consumer k is the probability $p_k(A)$ that consumer k chooses Product A. Likewise for Product B. With respect to the entire network, the market share of a Company is the sum, over all consumers in the network, of the probabilities that each consumer chooses their Product. The *bias* μ_k of consumer k is the difference in the probabilities of selecting A and B, i.e.,

$$\mu_k \stackrel{\Delta}{=} p_k(A) - p_k(B).$$

The *total bias* on the network is the sum of the biases of all consumers,

$$\begin{aligned} \sum_{k \in V} \mu_k &= \sum_{k \in V} [p_k(A) - p_k(B)] \\ &= \sum_{k \in V} p_k(A) - \sum_{k \in V} p_k(B) \ , \end{aligned} \tag{1}$$

the difference in market share [3] of the two companies. Note that Company A wants to *maximize*, whereas Company B seeks to *minimize*, total bias.

Companies A and B select respective *marketing allocations*, which are subsets of consumers to target with consumer-specific types of marketing, by learning models for the consumer choice probabilities $p_k(\cdot)$ that take into account the effect of marketing, by optimizing over expected total bias. It is this that we refer to as *A Marketing Game*.

The framework presented in this paper provides a firm foundation on which to construct a decision-influencing operation. However, in order for such an operation to be viable, it must leverage advances in so-called *affective computing* [8], which connects states of mind, for example preference between Products A and B, and the effects of such preference, for example content shared on social media. Moreover, in [33] we introduced a *marketing response* for each consumer that reflects consumers' individual responsiveness to a particular type of marketing. Such a marketing response will necessarily abide by the well-known Weber-Fecher or Stevens Laws [37] in quantifying the change in a consumer's perception of the value, or utility, of a Product, as a function of the marketing intensity, e.g., frequency and duration. In other words, this paper proposes a theoretical scaffold upon which to organize relevant social science contributions.

The following section discusses related work and the contributions of this paper. Section 3 provides overview on socially-contingent random utility and introduces the marketing-influenced parametrization thereof. Section 4 discusses the emphasis on *influences* in our model in the context of prior emphasis on so-called *influentials*. Section 5 briefly discusses the connection between consumer preferences and posts on social media. Section 6 outlines the basic data analytic components of *A Marketing Game*. Finally, Section 7 concludes with a discussion of limitations and future work.

2 Related Work

The problem of influencing decision-making on a social network has attracted a great deal of attention [15,12,35,19,45,27,1]. Some of this attention has focused on the role played by network topology [45,27]. Social networks have been characterized in different ways, the two most prominent being the *small-world* [44] and *scale-free* [2] properties. Small-world networks are defined as occupying an intermediate position between completely random and completely regular networks, in terms of so-called *clustering coefficient* and *average (shortest) path length*. Scale-free networks possess so-called power-law degree distributions, which result in a characteristic of relatively few *hubs*, individuals who are highly connected within the network, with most others connected to few others. There has been recent work showing that *spanning trees* of scale-free networks are typically themselves scale-free [21]. As such we feel that the so-called *star-chain* network illustrated in Figure 1 (a) is a useful atom to consider. The numerical analysis we present in Sections 6.1 and 6.3 will be with respect to this network.

In addition to network topology, there is the issue of *choice dynamics* on the network, the process by which consumers form and modify their preferences in response to *interactions* with neighboring consumers. Kempe et al [19] and Watts and Dodds [45] have examined dynamics ranging from contagion models inspired by epidemiology to threshold models tantamount to best-response in the parlance of so-called interaction games [4,27]. The language of contagion and epidemic can in part be traced to *The Tipping Point* [15] in which Gladwell draws analogy between widespread product adoption and the outbreak of disease. Unfortunately, the epidemic analogy is rather misleading with respect to product adoption. For example, if neighboring consumers i and j are "infected", respectively, with Products A and B, and i infects j, then j is no longer infected with B. But this is not how diseases work, and as such, epidemics is not a good model for the adoption of preferences on a network.

On the other hand, Gladwell also identified more germane forces influencing the spread of a Product. Through a number of examples he argues that *connectors*, *mavens*, and *salesmen* facilitate diffusion of product preference. In addition, he introduces the idea of a Product's *stickiness* to indicate the likelihood that consumers will continue to choose a product after gaining experience with it. A central argument of this paper is that these concepts, while informal, are nevertheless captured in a rather simple parametrization of consumer choice based on *random utility* [26]. In particular, the *socially-contingent* extension of random utility introduced by Blume [4], and explored by Montanari and Saberi [27] in the context of Product adoption, includes the *inherent bias* α_i of a consumer towards the Products, and *social biases* $\theta_{j\to i}$ and $\theta_{i\to j}$ indicating the influence that neighboring consumers i and j exert upon one another. That is to say, the inherent bias α_i captures the stickiness of Gladwell's formulation, while the social biases $\theta_{j\to i}$ and $\theta_{i\to j}$ capture the relative 'maven-ness' of neighboring consumers towards on another.

Montaneri and Saberi [27] considered the effect of network topology on preference adoption in the case of non-uniform inherent biases where all inherent biases favored the same Product. While we feel, and discuss briefly in Section 3.2, that such a model is useful for *certain* markets of social decision-making, it is ill-suited for modeling product adoption. For one, it, and indeed the majority of works that can be interpreted within the interaction game paradigm [35,19,45,14,9,1], model choice updates as best-responses to inherent and social biases, conditioned on the preferences of one's neighbors. Moreover and more importantly, the model in [27] does not include marketers for the Products. Indeed, the common approach of these works is *seeding* a Product at a select subset of consumers and then analyzing production adoption under best-response choice dynamics, without any ongoing efforts to market the Product.

In contradistinction, in our model consumers update their preferences *randomly* [1], according to the *logit* choice response model [26] with inherent biases

[1] It is important to note that in [4] and [27], among others, randomness is included in the choice updates. However, such *noisy* best-response dynamics are employed as a stratagem to force convergence to the payoff rather than risk dominant strategy.

$\{\alpha_i\}$ and social biases $\{\theta_{j\to i}\}$ determined from data. In [10] they consider the problem of *learning* the influences between neighboring consumers. Their learning criterion is a squared error metric with respect to a deterministic choice dynamics model. On the other hand, *random utility* [5,26] theory reflects the fact that, *with respect to a particular market*, consumer choices will *appear* random due to the fact that decisions made between alternatives A and B will nevertheless be influenced by considerations external to the market. In other words, utility is viewed as a parametrization of the *actual* frequencies with which consumers exhibit preference between alternatives. The modeler will decompose the parametrization into different influences that one suspects may be important, and which correspond to data that can be observed. That is, random utility theory is a truly data-driven framework for modeling consumer decision-making.

More important than the resulting stochastic choice dynamics, the power of the random utility parametrization is that by including *marketing* into the parametrization, Companies A and B can approach the problem of optimizing their *marketing allocations* within the framework of *reinforcement learning* [38]. That is, m^i_A and m^i_B are marketing biases applied to consumer i from Companies A and B, respectively. Decisions by Companies A and B as to which consumers to target with marketing will be determined by learning inherent and social biases, as well as the *marketing responses* [33] of individual consumers to different types of marketing. The marketing response and the level of investment by Company A, for instance, in marketing to consumer i, is what determines the value of the marketing bias m^i_A. The marketing responses will be learned by market research. On the other hand, inherent and social biases can be learned by well-known inference algorithms for graphical models [42]. Marketers are indicated in Figure 1 (b), and correspond to the *salesmen* of Gladwell's model.

Recently, Abebe et al have considered the problem of maximizing influence from the perspective of modifying consumers' so-called *susceptibiliites* [1]. Our argument here is that consumer i's susceptibility, if you will, can be decomposed into inherent bias α_i, social biases $\theta_{j\to i}$, and marketing biases m^i_A and m^i_B. In the context of this paper, then, modifying a consumer's susceptibility amounts to Company A, for example, investing more, presumably in *effective* marketing, to increase the marketing bias m^i_A applied to consumer i. In future work, it could correspond to efforts to influence social biases between neighboring consumers, for example the effect of rumor spreading.

This paper rigorously derives the marketing-influenced parametrization that enables incorporation of marketing responses into a model of consumer network decision-making. This paper likewise presents a general template for learning network biases from social media content. Such learned biases are fused with marketing responses gleaned from market research to form a model of network decision-making. We illustrate in Section 6.3 the importance of parameter estimation in optimizing marketing allocation. In particular, if Company A is unable to distinguish between an inherent bias in favor of Company B and marketing bias applied by Company B, what Company A determines to be the *best* allocation could in fact be the *worst*.

3 Choice Dynamics Model

Consumers choose between alternatives A and B in part as a result of the perceived utility of each alternative. Utility can be viewed as a scale for measuring differences in a consumer's perception of the respective value provided by each alternative [40,24]. This perceived value is partly due to objective matters such as price and monetary returns, but also to more subjective matters such as enhanced enjoyment and status. For example, prospect theory [20] posits that choices are often determined more by minimization of risk rather than maximization of monetary expectation, presumably because losing a gamble can result in a loss of status, and therefore greater utility is assigned to a more certain though less obviously beneficial possibility.

Random utility theory [5,26], on the other hand, is somewhat more agnostic, positing instead that consumers *maximize* utility, but that the utility assigned to an alternative by a consumer can be decomposed into *known* and *unknown* sources. Such a parametrization may, of course, obscure *explanation* of the observed frequencies of choice, for example, as provided by prospect theory [20]. In this section we introduce the marketer into the random utility parametrization of socially-contingent choice, which, as we discuss in Section 6, enables Companies to combine market research, data analytics, and simulation, to optimize marketing allocation.

Let

$$x_i = \begin{cases} 1 & \text{if consumer } i \text{ chooses } A \\ -1 & \text{if consumer } i \text{ chooses } B \end{cases} \tag{2}$$

numerically denote consumer i's *choice* or *preference*, and X_i the random variable associated with consumer i's *possible* choice. We will abuse notation and let x_i refer to both the numerical value (1 or -1) and the choice (A or B). Let $\mathbf{x} = (x_1, \ldots, x_{|V|})$ denote a *configuration* of choices on the network, and $\mathbf{X} = (X_1, \ldots, X_{|V|})$ the *random field* associated with choices on the network.

3.1 Random Utility Parametrization of Choice

In choosing between alternatives A and B, consumers seek to maximize between the utilities

$$U = \begin{bmatrix} u_A + \epsilon_A \\ u_B + \epsilon_B \end{bmatrix} , \tag{3}$$

where u_A and u_B are the known sources of utility assigned respectively to Products A and B, and ϵ_A and ϵ_B are the respective unknown sources of utility. For example, if a modeler opted to use expected monetary gain as the known sources of utility u_A and u_B, then a consumer's preference for a certain gain over a less certain but larger gain would be attributable to the *unknown* sources of utility ϵ_A and ϵ_B. In this case, by observing the *actual* frequencies with which consumers choose one alternative over the other, the modeler would fit a value for the parameter associated with numerical expectation that predicted monetary gains obtained from each of the alternatives.

When a consumer updates his choice by maximizing the utility in (3), he will choose Product A if $u_A + \epsilon_A > u_B + \epsilon_B$. Because the sources of utility ϵ_A and ϵ_B are unknown, we model them as random variables. Therefore, whether a consumer chooses Product A will likewise be random, with probability

$$p(u_A + \epsilon_A > u_B + \epsilon_B) = p(\epsilon_A - \epsilon_B > u_B - u_A) \ . \tag{4}$$

Assume that the unknown sources of utility ϵ_A and ϵ_B are distributed as the maxima of sequences of independent and identically distributed random variables. In this case (4) becomes a *logit* response distribution [26,39], i.e.,

$$p(A) = \frac{e^{u_A}}{e^{u_A} + e^{u_B}} \ , \tag{5}$$

and likewise for $p(B)$. As such, the utilities u_A and u_B can be viewed as a parametrization of individual choice probabilities. In part for this reason, the logit model is common in marketing research [18,25]. One can also derive randomness in individual choice by hypothesizing *bounded rationality* [36] on the part of consumers.

Different assumptions about ϵ_A and ϵ_B will lead to different choice rules. For example, if the unknown sources of utility for different alternatives are instead modeled as dependent, one instead derives a *nested* logit model, which can be viewed as an iterative logit model akin to Tversky's model of aspect elimination [41]. Moreover, if ϵ_A and ϵ_B are modeled as *sums* rather than *maxima* of i.i.d. variables, one derives a *probit* rather than a *logit* model [39].

3.2 Socially-Contingent Parametrization of Choice

When the parametrization of consumers' utilities are contingent upon the choices of other consumers, the interdependence of utility is referred to as a *game* [4]. Such socially contingent choice can be extended to networks of consumers [4,7] whose individual choices are contingent upon a small subset of other consumers, referred to as *neighbors*. Let ∂i denote the set of neighbors of consumer i. The known sources of utility u_A and u_B for consumer i will be decomposed additively into utility derived through agreement or disagreement with his neighbors in ∂i.

Let $u^i_{A|A}$ and $u^i_{B|A}$ be the respective known utilities that consumer i derives from Products A and B conditioned on neighbor $j \in \partial i$ having chosen Product A. Likewise for $u^i_{A|B}$ and $u^i_{B|B}$ conditioned on j having chosen B. This dependence can be summarized by the matrix

$$u^i_{|j} = \begin{bmatrix} u^i_{A|A} & u^i_{A|B} \\ u^i_{B|A} & u^i_{B|B} \end{bmatrix} ,$$

where the column corresponds to the choice by j. Following Blume [4], we can re-write $u^i_{|j}$ as

$$u^i_{|j} = \begin{bmatrix} \theta_{j\to i} & -\theta_{j\to i} \\ -\theta_{j\to i} & \theta_{j\to i} \end{bmatrix} + \begin{bmatrix} \alpha_{i|j} & \alpha_{i|j} \\ -\alpha_{i|j} & -\alpha_{i|j} \end{bmatrix} \ , \tag{6}$$

where

$$\theta_{j \to i} \quad = \quad \frac{u^i_{A|A} + u^i_{A|B} - u^i_{B|A} - u^i_{B|B}}{4}$$

and

$$\alpha_{i|j} \quad = \quad \frac{u^i_{A|A} - u^i_{A|B} - u^i_{B|A} + u^i_{B|B}}{4} \ .$$

Here, $\theta_{j \to i}$ is the *social bias* exerted upon consumer i by neighbor j, due to whether i makes the same choice as j. We define

$$\begin{aligned} \alpha_i &\triangleq \sum_{j \in \partial i} \begin{bmatrix} \alpha_{i|j} \\ -\alpha_{i|j} \end{bmatrix} \\ &= \begin{bmatrix} \alpha_i \\ -\alpha_i \end{bmatrix} \end{aligned} \tag{7}$$

to be the *inherent bias* of consumer i, with $\alpha_i > 0$ indicating a bias in favor of Product A and $\alpha_i < 0$ a bias in favor of Product B.

Using the numerical representation of choice (2), the utility vector (3) for consumer i, conditioned on the choices $\mathbf{x}_{\partial i}$ of his neighbors, can now be decomposed as

$$U_{|\mathbf{x}_{\partial i}} = \begin{bmatrix} \alpha_i + \sum\limits_{j \in \partial i} \theta_{j \to i} x_j + \epsilon^i_A \\ -\alpha_i - \sum\limits_{j \in \partial i} \theta_{j \to i} x_j + \epsilon^i_B \end{bmatrix} \ . \tag{8}$$

Applying the socially-contingent decomposition of utility (8) to the random utility update (5) yields the following choice dynamics:

$$p(x_i | \mathbf{x}^{(t)}_{\partial i}) = \frac{\exp\left\{ \sum\limits_{j \in \partial i} \theta_{j \to i} x_i x^{(t)}_j + \alpha_i x_i \right\}}{Z_{i|\mathbf{x}^{(t)}_{\partial i}}} \ , \tag{9}$$

referred to as *Glauber dynamics* [16] in the statistical mechanics literature, where $Z_{i|\mathbf{x}^{(t)}_{\partial i}}$ is the normalizing constant referred to as the (local) partition function at i conditioned on $\mathbf{x}_{\partial i}$ at time t.

Best- or near best-response dynamics [4,46,13,28,45,27,14] have long been considered . As discussed in [39], scaling the utilities u_A and u_B by a constant $\beta > 0$ amounts to adjusting the relative importance of the known and unknown sources of utility. For example, the probability of a consumer choosing Product A becomes

$$p(\beta u_A + \epsilon_A > \beta u_B + \epsilon_B) \quad = \quad p(u_A + \frac{\epsilon_A}{\beta} > u_B + \frac{\epsilon_B}{\beta}) \ ,$$

in which case the limit $\beta \to \infty$ would amount to a market in which consumers' perception of utility for Products A and B significantly outweigh sources of utility external to the market; moreover, a market in which the modeler is fully aware of the within-market sources of utility. Such a construction, considered explicitly in [27] and implicitly in [45] and [14], is useful from the *scientific* objective of understanding social norms, where emphasis on *fitting in* likely outweighs influences from other markets. However, from the *engineering* objective of using data to influence consumer decision-making, consumer choice should not be modeled under a $\beta \to \infty$ scaling of known utility, but rather stochastically, as (9), where utilities α_i and $\theta_{j\to i}$ can be estimated from data as those that predict actual choice frequencies.

More importantly, one needs to include the marketer into the model so that companies can account for the influence that marketing has on choice dynamics.

3.3 Including the Marketer in Choice Parametrization

On a purely mathematical level, the primary contribution of this paper is incorporating the influence of marketing for Products A and B into the above parametrization of choice dynamics. In pioneering work, Harold Laswell [22] examinined the social role that media advertising plays in consumer decision-making. Around the same time, David Ogilvy [29] opened what would become *Ogilvy and Mather*, ushering in an age of *content marketing*, in which emphasis was placed on communicating with consumers in a more informative rather than promotional manner. It therefore makes sense to view the marketer as a social connection with a constant preference.

We include in the socially-contingent decomposition of utility (8) the *marketing biases* $m_A^i > 0$ and $m_B^i > 0$ applied by Companies A and B, respectively, to consumer i, as

$$U_{|\mathbf{x}_{\partial i}} = \begin{bmatrix} \alpha_i + m_A^i - m_B^i + \sum_{j\in\partial i} \theta_{j\to i} x_j + \epsilon_A^i \\ -\alpha_i - m_A^i + m_B^i - \sum_{j\in\partial i} \theta_{j\to i} x_j + \epsilon_B^i \end{bmatrix} . \tag{10}$$

The marketing biases applied to a consumer are tantamount to the social biases neighboring consumers exert upon each other, for example as indicated in Figure 1 (b). The strength of the marketing biases will depend on the *marketing response* [33] of consumer i to the particular type of marketing, which indicates the degree of influence as a function of marketing intensity. This is discussed briefly in Section 6.2.

The choice dynamics now become

$$p(x_i|\mathbf{x}_{\partial i}^{(t)}) \;=\; \frac{\exp\left\{ \sum_{j\in\partial i} \theta_{j\to i} x_i x_j^{(t)} + \theta_i x_i \right\}}{Z_{i|\mathbf{x}_{\partial i}^{(t)}}} , \tag{11}$$

where $\theta_i = \alpha_i + m_A^i - m_B^i$ is the *direct bias* of consumer i.

4 Influences vs. Influentials

Slightly before Lasswell's work, Paul Lazarsfield et al analyzed the influence of media on consumer choice in a presidential election, concluding that for most consumers, social bias was more influential than the applied bias of media marketing. Based on these findings, they introduced the *two-step model of communication* [23] whereby the role of a marketer was to target *influential* opinion leaders, as they were the gateway to everyone else. In the *The Tipping Point* [15], Malcolm Gladwell revisited the idea of *influentials* with his *salesmen*, *mavens*, and *connectors*.

In contrast, it is important to note that our model does not stress influential *individuals*, but rather *influences*. For example, if $\theta_{i\to j} > \theta_{j\to i}$, for each neighbor $j \in \partial i$, then we might refer to consumer i as a *maven*, one who exerts a strong influence over the preferences of others. On the other hand, we might simply say that i is a maven in his interactions with neighbor j_1, if $\theta_{i\to j_1} > \theta_{j_1\to i}$, even if $\theta_{i\to j_2} < \theta_{j_2\to i}$, for some other neighbor $j_2 \in \partial i$. If i is a hub in the network, or connects otherwise distant regions of the network, we might refer to consumer i as a *connector*. Moreover, an inherent bias α_i can be interpreted as the difference in *stickiness* between Product A and Product B, with respect to consumer i.

The relative importance of these influences will depend on the overall constellation of influences. Indeed, Watts and Dodds [45] found that under certain combinations of inherent and social biases[2], targeting connectors would result in a so-called *cascade*, while with other combinations, targeting connectors did not lead to a cascade. Our objective is not to affirm or contradict the so-called *influentials hypothesis*, nor to identify *a priori* which *influences* are more important. Rather, the relative importance of influences will manifest in simulation of the network under influences estimated from data.

As indicated above, our primary mathematical contribution is including the influence of the *salesmen*, through marketing biases m^i_A and m^i_B.

5 Social Media Posts

At a given time t, the configuration of choices $\mathbf{x}^{(t)} = (x_1^{(t)}, \ldots, x_{|V|}^{(t)})$ on the network represent preferences for the two Products. There is a corresponding configuration $\mathbf{y}^{(t)} = (y_1^{(t)}, \ldots, y_{|V|}^{(t)})$ of consumer posts. Posts can be an image, a block of text, or a combination of the two. If at time t consumer i prefers Product A, the post $y_i^{(t)}$ could reflect *positive sentiment* with respect to A or *negative sentiment* with respect to B [30].

When consumer i updates his choice to $x_i^{(t+1)}$, he does so based on his inherent bias towards Products A and B, the applied bias from Companies A and B, the social biases from his neighbors, and his understanding of the preferences indicated in his neighbors' posts $\{y_j^{(t)} : j \in \partial i\}$. We assume in this paper that

[2] They referred to the *susceptibility* of a consumer.

the post $y_i^{(t)}$ is perfectly correlated with consumer i's choice $x_i^{(t)}$, so that consumer i's neighbors can be said to observe i's choice $x_i^{(t)}$. In general, however, there may be some ambiguity between consumer i's post $y_i^{(t+1)}$ and his actual preference $x_i^{(t+1)}$ [31]. When a consumer updates his choice to $x_i^{(t+1)}$, he creates a new post $y_i^{(t+1)}$.

In Section 6 we briefly discuss the use of deep learning [17] and sentiment analysis [8] to detect semantic relationships between objects and topics in consumers' social media posts. In order to fully leverage these tools, Companies will additionally need models correlating a user's preference towards Products (commercial or electoral) and the semantic content of their social media posts. For example, consumers i and j may both prefer political candidate A, but whereas consumer i may post favorable content with respect to candidate A, consumer j may post unfavorable content with respect to candidate B.

6 Analytics of *A Marketing Game*

As mentioned in Section 1, introducing marketing biases into the parametrization of choice dynamics places *A Marketing Game* within the purview of *reinforcement learning*. That is, Companies A and B will *learn* marketing biases m_A^i and m_B^i, respectively, through market research; direct biases $\theta_i = \alpha_i + m_A^i - m_B^i$ and social biases $\theta_{j\to i}$ through a combination of *deep learning* [17] and graphical model inference algorithms [42,32]; and use the resulting learned model of network choice dynamics to select the *marketing allocation* that optimizes predicted market share.

To be sure, successfully implementing this approach will require considerable interdisciplinary effort. Namely: on the marketing side, we need to understand how users' mental states will *respond to* content, so that we can create content *marketing*; on the analytics side, we need to understand how users' mental states *create* content, so that we can respond with appropriate content *analytics*; in particular, we need to understand how to use *deep learning* to infer from posted social media content consumer preferences that are in turn influenced by marketing content posted by Companies.

In this section we briefly discuss the basic components of an analytics pipeline, illustrated in Figure 2, for leveraging the marketing-influenced socially-contingent parametrization of consumer choice introduced in Section 3.3. The two components that interact directly with the network are the *API / data collection* module, which scrapes user posts from social media, and the *content marketing* module, which exposes users to marketing content analogous to the social media posts of a user's neighbors. In Section 6.1 we discuss estimation of direct and social biases from social media posts. In Section 6.2 we discuss marketing research required to estimate consumers' marketing responses. In Section 6.3 we discuss simulation of the network under candidate marketing allocations and illustrate the importance of distinguishing between inherent and marketing biases in favor of the other Company.

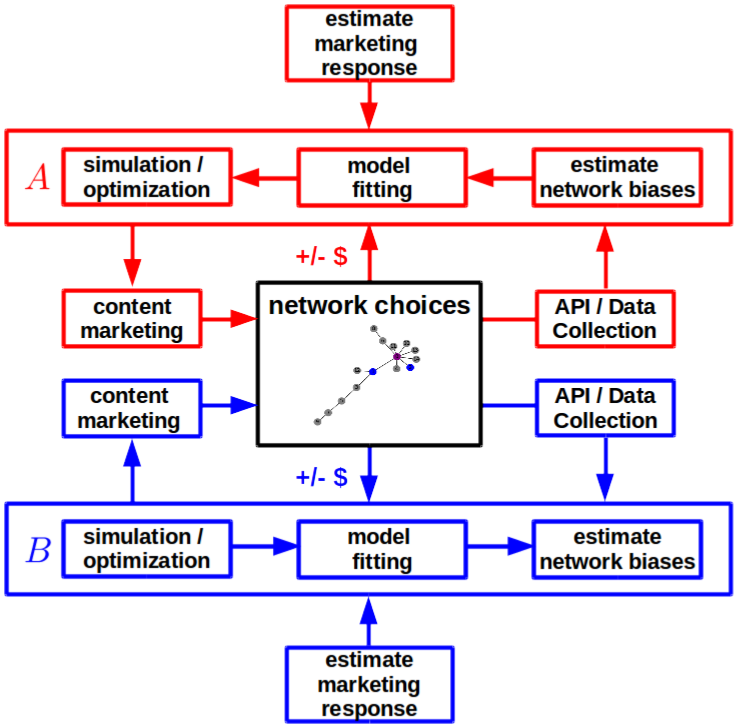

Fig. 2: Block diagram illustrating data analytic components of *A Marketing Game*.

6.1 Estimation of Direct and Social Biases

A Company will use an *application programming interface* (API) to collect data, for example *posts* $\mathbf{y}^{(t-1)}, \ldots, \mathbf{y}^{(t-T)}$, from a social media network. For a given consumer i and post $y_i^{(t-\tau)}$, Company A uses machine learning algorithms, for example a *recurrent* or *convolutional* neural network [17] to determine the subject of the post, and *sentiment analysis* [8] to determine the consumers' attitudes with respect to the subject.

Using models correlating preferences and posts, Company A will form an estimate $\hat{x}_i^{(t-\tau)}$ of consumer i's preference, which can be viewed as a *noisy* version of $x_i^{(t-\tau)}$, the noise being determined by the accuracy of the machine learning algorithms. From the perspective of Company A, the preferences ($\mathbf{X}$) constitute a *hidden* Markov random field in which noisy observations $\{\hat{\mathbf{x}}^{(t-\tau)}\}$ are observed, while true choices $\{\mathbf{x}^{(t-\tau)}\}$ are unobserved.

In order to simplify things, assume that Company A's machine learning algorithms are perfect, so that Company A observes the true sequence of choice configurations $\{\mathbf{x}^{(t-\tau)}\}$. Company A can leverage the Markov property of $\mathbf{X}$, that knowing the choices of a consumer's neighbors renders the consumer's choices in-

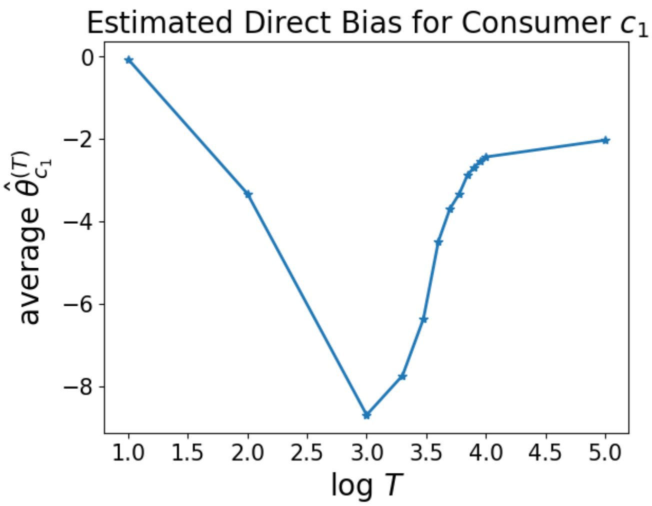

Fig. 3: Estimates $\hat{\theta}_{c_1}$ of the direct bias for hub consumer c_1 in Figure 1 (a) as a function of sample complexity. Note that the estimates converge to a direct bias of $\hat{\theta}_{c_1} = -2$, which favors Company B. Note also that this estimate does not disambiguate $\alpha_{c_1} + m_A^{c_1} - m_B^{c_1}$. In Sections 6.2 and 6.3 we will assume that this estimation is performed by Company A and used in the evaluation of candidate marketing allocations.

dependent of choices outside of his neighborhood, by minimizing the *conditional description length* [32]

$$\bar{D}(x_i^{(t-T:t)}|\mathbf{x}_{\partial i}^{(t-T:t)};\theta_{\bar{i}}) = -\sum_{\tau=1}^{T} \log p(x_i^{(t-\tau)}|\mathbf{x}_{\partial i}^{(t-\tau)};\theta_{\bar{i}})$$

of consumer i's choices *conditioned* on the choices of his neighbors. Here, $x_i^{(t-T:t)}$ indicates the sequence of observations $x_i^{(t-T)}, \ldots, x_i^{(t)}$, and $\theta_{\bar{i}} \triangleq \theta_i \cup \{\theta_{j\to i} : j \in \partial i\}$ denotes the parameter specifying consumer i's choice. Figure 3 illustrates estimates of the direct bias for the hub consumer c_1 in Figure 1 (a). When the machine learning algorithm of Company A is not perfect and estimated parameters $\tilde{\theta} = (\tilde{\theta}_i, \tilde{\theta}_{j\to i})$ must be determined using noisy observations $(\hat{\mathbf{x}}^{(t-\tau)})$, Company A can use a variant of the well-known *expectation-minimization* (EM) algorithm [11].

The reason for focusing on estimating consumer preferences rather than simply social media posts is that our ultimate concern is whether a consumer will purchase a product or vote for a candidate. Therefore, we want to go from *posts* back to states of mind, i.e., *preferences* with respect to Product alternatives. This implies that one has to be somewhat judicious regarding *what* posts one "pulls" for the purposes of estimating network biases.

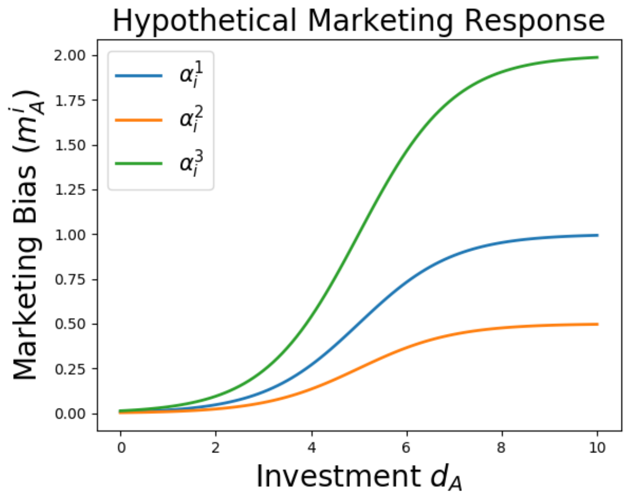

Fig. 4: Hypothetical marketing responses indicating marketing bias m_A^i applied to consumer i by Company A as a function of investment d_A^i, for different values of inherent bias α_i.

6.2 Estimating Marketing Response

The relationship between marketing investment in a consumer and the resulting perceived utilities m_A^i and m_B^i would, in practice, be determined by market research. Such a relationship between stimulus (i.e., marketing) intensity and perception of value will likely obey the well-known Weber-Fechner [6] or Stevens [37] Laws.

For example, there would be a saturation effect where additional investment has only negligible influence on consumer choice. Moreover, the response of a consumer to marketing by a Company would likely depend on any inherent bias of the consumer towards Products A or B. For example, if a consumer has a bias towards one or the other product, we would expect that he will be less responsive to marketing from *both* Companies than if he has no bias. For instance, if a consumer is biased in favor of Product A, marketing by Company A will only incrementally add to the effective attractiveness of Product A. On the other hand, if a consumer is biased in favor of Product B, marketing by Company B can only do so much to attract the consumer. Figure 4 illustrates hypothetical marketing responses for a consumer.

In the previous section, Company A forms an estimate

$$\tilde{\theta}_{c_1} = \tilde{\alpha}_{c_1} + \tilde{m}_A^{c_1} - \tilde{m}_B^{c_1} \tag{12}$$

of the direct bias at consumer c_1. If Company A knows the shape of consumer c_1's marketing response, for example as illustrated in Figure 4, then given an estimate

$\tilde{\alpha}_{c_1}$ of consumer c_1's inherent bias, Company A can form an estimate $\tilde{m}_A^{c_1}$ of the marketing strength applied to consumer c_1 at a given level of investment. Company A can then estimate $\tilde{m}_B^{c_1}$ from (12). In particular, it can determine whether a direct bias at consumer c_1 in favor of Product B is due to inherent bias or marketing. We will see in the next section that being able to disambiguate inherent and applied bias can have dramatic consequences for the allocation of marketing resources. Estimating inherent biases and marketing responses can be carried out with surveys, focus groups, and A/B testing.

6.3 Simulation and Optimization

The estimated direct and social biases from Section 6.1 will be combined with the estimated marketing responses from Section 6.2 into a model of network decision-making for candidate marketing allocations. Company A will in general simulate the corresponding choice dynamics (11) and select the allocation that optimizes expected market share. To illustrate, we consider a simplified scenario: symmetric social biases on the star-chain network of Figure 1 (a), where all consumers except the hub consumer c_1 have inherent bias $\alpha_i = 0$ and marketing biases $m_B^i = 0$ from Company B. We consider optimization of Company A's allocation under three cases for the direct bias of c_1. The first two correspond to the estimated direct bias $\hat{\theta}_{c_1} = -2$ shown in Figure 3. In other words, how does Company A's optimal allocation depend on whether that direct bias is an inherent bias in favor of Company B versus a marketing bias from Company B? The third case considers that consumer c_1 has an inherent bias in favor of Company A.

Blume [4] shows that if the social biases are symmetric, that is, $\theta_{j \to i} = \theta_{i \to j} = \theta_{ij}$, for all pairs of neighboring consumers i and j, then the dynamics of (11) converge to the equilibrium *Gibbs* distribution given by

$$p(\mathbf{x};\theta) = \frac{1}{Z(\theta)} \exp\{\sum_{\{i,j\}} \theta_{ij} x_i x_j + \sum_{i \in V} \theta_i x_i\} \ , \qquad (13)$$

where

$$Z(\theta) = \sum_{\mathbf{x}} \exp\{\sum_{\{i,j\}} \theta_{ij} x_i x_j + \sum_{i \in V} \theta_i x_i\}$$

is the (global) partition function. The total bias (1) can be computed as

$$\sum_{k \in V} \mu_k = \sum_{k \in V} \frac{Z_k(A) - Z_k(B)}{Z_k(A) + Z_k(B)}, \qquad (14)$$

where the vector Z_k, with components

$$Z_k(A) = \sum_{\mathbf{x}: x_k = A} \exp\{\sum_{\{i,j\} \in E} \theta_{ij} x_i x_j + \sum_{i \in V} \theta_i x_i\}$$

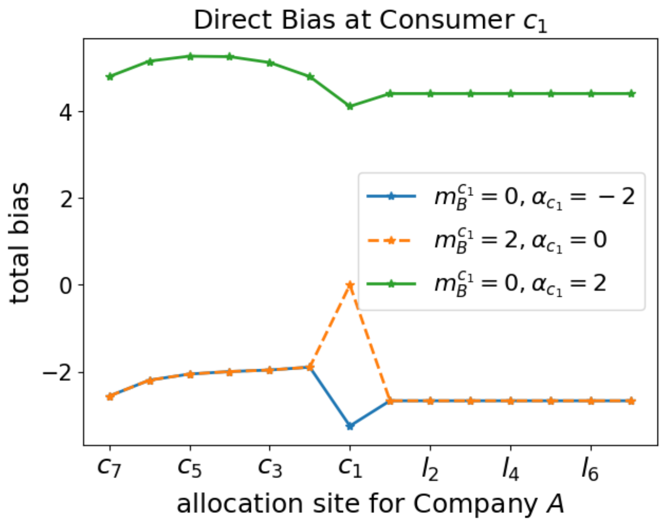

Fig. 5: Total bias on star-chain network with a direct bias at consumer c_1. The direct bias is either an inherent bias in favor of Product B, i.e., $m_B^{c_1} = 0$ and $\alpha_{c_1} = -2$; an applied marketing bias from Company B, i.e., $m_B^{c_1} = 2$ and $\alpha_{c_1} = 0$; or an inherent bias in favor of Product A, i.e., $m_B^{c_1} = 0$ and $\alpha_{c_1} = 2$.

and

$$Z_k(B) = \sum_{\mathbf{x}:x_k=B} \exp\{\sum_{\{i,j\}\in E} \theta_{ij} x_i x_j + \sum_{i\in V} \theta_i x_i\} ,$$

is the *belief* for consumer k.

For the star-chain network, we can compute $Z_k(A)$ and $Z_k(B)$ using *Belief Propagation* [42]. Figure 5 illustrates the total bias that results from Company A allocating a marketing bias m_A^i to a single consumer in this star-chain network, for the three different cases of direct bias at the hub consumer c_1. We see that if Company A knows only that there is a direct bias in favor of B, without knowing whether that bias is due to inherent bias or the effect of marketing, then Company A could actually make the *worst* marketing allocation when it *thinks* it is making the best. That is, Company A placing a unit of marketing allocation at the hub c_1 is optimal if the direct bias at c_1 in favor of B is due to marketing, but is the worst allocation if it is an inherent bias in favor of B. This comports with the intuition that marketing to consumers who are loyal towards an opposing brand is not a good investment. On the other hand, if the direct bias at c_1 is an inherent bias in favor of Company A, then Company A allocating its single unit of marketing to c_1 likewise results in the lowest possible total bias. The reason is that consumer c_1's inherent bias towards Product A already serves as a form of marketing to his neighbors, and thus greater total bias can be achieved by Company A allocating elsewhere.

7 Summary, Limitations, and Future Directions

We have derived the marketing-influenced parametrization of socially-contingent consumer choice, discussed the estimation of network biases from data, and analyzed a star-chain network to illustrate optimization of marketing allocation.

Limitations of this paper mainly revolve around work that remains to be done in order to implement this program on a large scale. For example, looking at real data from a social network, and implementing machine learning algorithms to infer consumer preferences. This will inevitably require us to understand asynchronous choice dynamics, as users post content at varying rates. Moreover, while the simplified scenario in this paper consisted of symmetric social biases, which in the case of two Products, leads to a tractable equilibrium Gibbs distribution, in general social biases will not be symmetric and the resulting dynamics will not have an equilibrium. For instance, social connections on Twitter are asymmetric. Preliminary analysis with asymmetric social biases [34] suggests that, while not possessing an *equilibrium* per se, such asymmetric dynamics do have as a stationary distribution the Gibbs equilibrium corresponding to symmetric dynamics, albeit with different direct biases. This would be significant regarding actual implementation of this model, as it has been shown [43] that suboptimal variational methods with respect to an equilibrium Gibbs model can yield better performance than Monte Carlo simulation when computing resources are at a premium.

Furthermore, as highlighted in [10], it is important to consider the case where the choice dynamics themselves are non-stationary. Social connections and biases change over time, and it is unlikely that a given set of network biases will remain constant long enough for a stationary distribution to be attained. Again, the work in [34] makes some inroads into consideration of transient phase analysis.

Lastly, there is considerable need to focus on market research to understand marketing responses and the use of such information to disambiguate estimated direct biases.

References

1. R. Abebe, J. Kleinberg, D. Parkes, and C.E. Tsourakakis. *Opinion Dynamics with Varying Susceptibility to Persuasion.* In KDD 2018: The 24th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD). August 2018. London, UK.
2. A-L Barabasi and E. Bonabeau, “Scale-Free Networks,” *Scientific American*, May 2003, pp. 60-69.
3. D.E. Bell, R.L. Keeney, and J.D.C. Little, “A Market Share Theorem,” *Journal of Marketing Research*, Vol. 12, No. 2, May 1975.
4. L.E. Blume, “Statistical Mechanics of Strategic Interaction,” *Games and Economic Behavior*, Vol. 5, Issue 3, 1993.
5. H.D. Block and J. Marschak, *Random orderings and stochastic theories of responses*, In Contributions to Prob. and Stats., Stanford University Press, 1960.
6. S.H. Britt, “How Weber’s Law Can Be Applied to Marketing,” *Business Horizons*, **18**(1), 1975.

7. W.A. Brock and Durflauf, “Discrete Choice with Social Interactions,” *Review of Economic Studies*, Vol. 68, 2001.
8. E. Cambria. *Affective Computing and Sentiment Analysis.* IEEE Intelligent Systems. Vol. 31, No. 2, 2016.
9. F. Chierichetti, J. Kleinberg, and S. Oren. *On discrete preferences and coordination.* In ACM Conf. on Electonic Commerce, 2013, pp. 233-250.
10. A. De, S. Bhattacharya, P. Bhattacharya, N. Ganguly, and S. Chakrabarti. *Learning a Linear Influence Model from Transient Opinion Dynamics.* The 23rd ACM Intl. Conf. on Info. Mngmt. ACM, 2014, pp. 401-410.
11. A.P. Dempster, N.M. Laird, and D.B. Rubin, *Max. Likelihood from Incomplete Data via the EM Algorithm*, Jrnl. Royal Stat. Soc., Series B, Vol. 39, No. 1, 1977.
12. P. Domingos and M. Richardson. *Mining the network value of customers.* In Proceedings of SIGKDD 2001, San Franciso, CA, pp. 57-66.
13. G. Ellison, *Learning, Local Interaction, and Coordination*, Econometrica, Vol. 61, No. 5, September 1993.
14. A. Fazeli and A. Jadbabaie, *Game Theoretic Analysis of a Strategic Model of Competitive Contagion and Product Adoption in Social Networks*, http://repository.upenn.edu/ese_papers/618, December 2012.
15. M. Gladwell, *The Tipping Point*, Little, Brown, and Company, 2000.
16. R.J. Glauber, *Time-Dependent Statistics of the Ising Model*, Journal of Mathematical Physics, Vol. 4, 1963.
17. I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning.* MIT Press, Cambridge, MA, 2016.
18. P.E. Green, F.J. Carmone, and D.P. Wachspress, *On the Analysis of Qualitative Data in Marketing Research*, Journal of Marketing Research, February 1977.
19. D. Kempe, J. Kleinberg, and E. Tardos. *Influential nodes in a diffusion model for social networks.* In Proceedings of the 32nd Intl. Conf. on Automata, Languages, and Programming (ICALP) 2005, 1127-1138.
20. D. Kahneman and A. Tversky, *Prospect Theory: An Analysis of Decision under Risk*, Econometrica, Vol. 47, No. 2, March 1979.
21. D-H Kim, J.D. Noh, and H. Jeong, *Scale-free trees: The skeletons of complex networks*, Physical Review E, Vol. 70, 2004.
22. H.D. Lasswell, *The structure and function of communication in society*, Ed. L. Bryson, The Communication of Ideas, Harper and Brothers, 1948.
23. P.F. Lazarsfield, B. Berelson, and H. Gaudet, *The People’s Choice: how the voter makes up his mind in a presidential campaign*, Columbia University Press, 1944.
24. D. Luce, *Individual Choice Behavior*, Dover, 1959.
25. N.K. Malhotra, *The Use of Linear Logit Models in Marketing Research*, Journal of Marketing Research, February 1984.
26. D. McFadden, *Conditional Logit Analysis of Qualitative Choice Behavior*, In Frontiers in Econometrics, New York: Academic Press, 1974.
27. A. Montanari and A. Saberi, *The spread of innovations in social networks*, PNAS, Vol. 107, No. 47, 2010.
28. S. Morris, *Contagion*, Review of Economic Studies, Vol. 67, 2000.
29. D. Ogilvy, *Ogilvy On Advertising*, Vintage, 1985.
30. K. and V. Ravi. *A survey on opinion mining and sentiment analysis: tasks, approaches and applications.* Knowledge-Based Systems. Vol. 89, Nov. 2015, pp. 14-46.
31. A. Reyes, R. Rosso, and D. Buscaldi, *From humor recognition to irony detection: The figurative language of social media*, Data & Knowledge Eng., 2012, Vol. 74.

32. M.G. Reyes and D.L. Neuhoff. Minimum Conditional Description Length estimation of Markov random fields. In Information Theory and Applications Workshop. February 2016.
33. M.G. Reyes. *A Marketing Game: a rigorous model for strategic resource allocation.* In ACM Workshop on Machine Learning in Graphs, London, UK, Aug. 2018.
34. M.G. Reyes. *Stability of a Directed Ising Model with Non-Uniform Magnetic Field.* To appear in "Pattern Formation and Stability in Magnetohydrodynamics", Ed. N.O. Rojas, IntechOpen, London, UK.
35. M. Richardson and P. Domingos. *Mining the network value of customers.* In Proceedings of SIGKDD 2002, Edmonton, Alberta, Canada.
36. H.A. Simon, *A Behavioral Model of Rational Choice*, The Quarterly Journal of Economics, Vol. 69, No. 1, 1955.
37. S.S. Stevens. *To Honor Fechner and Repeal His Law.* Science, Vol. 133, No. 3446, 1961.
38. R.S. Sutton and A.G. Barto. *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA, 1998.
39. K. Train, *Discrete Choice Models with Simulation*, Cambridge Univ. Press, 2002.
40. L.L. Thurstone, *Psychological Analysis*, The Amer. Jrnl. Psych., Vol. 38, No. 3, July 1927.
41. A. Tversky, *Elimination by Aspects: A Theory of Choice*, Psychological Review, Vol. 79, No. 4, July 1972.
42. M. Wainwright and M. Jordan, *Graphical Models, Exponential Families, and Variational Inference*, Technical Report, UC Berkeley, 2003.
43. Wainwright MJ. *Estimating the "Wrong" Graphical Model: Benefits in the Computation-Limited Setting.* Jrnl. of Machine Learning Research. 2006, Vol. 7.
44. D.J. Watts and S.H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature, Vol. 393, June 4, 1998.
45. D.J. Watts and P.S. Dodds , *Influentials, Networks, and Public Opinion Formation*, Jrnl. Consumer Research, Vol. 34, December 2007.
46. H.P. Young, *The Evolution of Conventions*, Econometrica, Vol. 61, No. 1, 1993.